engineering proceedings

MDPI

Proceeding Paper

Forecasting of Signals by Forecasting Linear Recurrence Relations †

Nina Golyandina * and Egor Shapoval

Faculty of Mathematics and Mechanics, St. Petersburg State University, 199034 St. Petersburg, Russia; st063753@student.spbu.ru

* Correspondence: n.golyandina@spbu.ru

† Presented at the 9th International Conference on Time Series and Forecasting, Gran Canaria, Spain, 12–14 July 2023.

Abstract: The forecasting of a signal that locally satisfies linear recurrence relations (LRRs) with slowly changing coefficients is considered. A method that estimates the local LRRs using the subspace-based method, predicts their coefficients and constructs a forecast using the LRR with the predicted coefficients is proposed. This method is implemented for time series that have the form of a noisy sum of sine waves with modulated frequencies. Linear and sinusoidal frequency modulations are considered. The application of the algorithm is demonstrated with numerical examples.

Keywords: time series; signal; forecasting; singular spectrum analysis; linear recurrence relation

check for updates

Citation: Golyandina, N.; Shapoval, E. Forecasting of Signals by Forecasting Linear Recurrence Relations. *Eng. Proc.* **2023**, *39*, 12. https://doi.org/10.3390/engproc2023039012

Academic Editors: Ignacio Rojas, Hector Pomares, Luis Javier Herrera, Fernando Rojas and Olga Valenzuela

Published: 28 June 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Let us consider the problem of forecasting time series using singular spectrum analysis (SSA) [1–6]. The theory of SSA is quite well-developed; there are many papers with applications of SSA to real-life time series (see, for example, [6] and [7] (Section 1.7) with short reviews). SSA does not require a time-series model to construct the decomposition into interpretable components such as the trend, periodicities, and noise. However, for prediction in SSA, it is assumed that the signal (the deterministic component of the time series) $\mathsf{S}_N = (s_1, \dots, s_N)$ satisfies some model, in particular, a linear recurrence relation (LRR) with constant coefficients (maybe, approximately):

$$s_{i+d} = \sum_{k=1}^{d} a_k s_{i+d-k} \quad \text{for } i = 1, \dots, N-d. \tag{1}$$

This assumption is valid for signals in the form of a sum of products of polynomial, exponential, and sinusoidal time series, in particular, a sum of exponentially modulated periodic components. SSA also works for the case of trend extraction and the general case of amplitude-modulated harmonics, where the model is satisfied approximately. However, SSA is not applicable if the signal locally satisfies a changing LRR. An example of such a signal is sinusoidal frequency-modulated time series. This paper aims to construct a method for the prediction of time series locally governed by changing LRRs, staying within the framework of SSA.

Let us consider the model of time series in the form of a noisy signal, where the signal is locally governed by LRRs with slowly time-varying coefficients. A local version of SSA has already been considered earlier for signal estimation [8]. However, it results in different approximations of the segments of the time series and the prediction can be performed based on the last segment only. In this paper, a local modification of the recurrent SSA prediction, based on the construction of a prediction of the coefficients of the local LRRs, is proposed. This modification was applied to time series in which the signal is a sum of sinusoids with time-varying frequencies having non-intersecting frequency ranges, where the instantaneous frequency of each summand is slowly varying.

2. Basic Notions

2.1. Linear Recurrence Relations

Let us introduce several definitions. By time series of length N we mean a sequence of real numbers $\mathsf{X}_N = (x_1, \dots, x_N) \in \mathbb{R}^N$. Consider a time series in the form of the sum of a signal and noise $\mathsf{X}_N = \mathsf{S}_N + \mathsf{R}_N$ and state the problem of signal forecasting.

A time series S_N satisfies a linear recurrence relation (LRR) of order d if there exists a sequence $\{a_i\}_{i=1}^d$ such that $a_d \neq 0$ and (1) takes place.

A time series governed by an LRR satisfies a set of LRRs, one of which has minimal order; we will call it minimal LRR. Among the set of the governing LRRs, there is the so-called min-norm LRR with the minimum norm of coefficients, which suppresses noise in the best way [5] and is used for the recurrent SSA forecasting.

The characteristic polynomial of LRR (1) is defined as

$$P_d(\mu) = \mu^d - \sum_{k=1}^{d} a_k \mu^{d-k}.$$

The roots of the characteristic polynomial of the minimal LRR are called signal roots. The characteristic polynomial corresponding to an LRR governing the signal includes the signal roots, among others.

Remark 1. *Since* $P_m(\mu) = \prod\limits_{k=1}^{m} (\mu - \mu_k) = \mu^m - \sum\limits_{k=1}^{m} a_k \mu^{m-k}$, *the roots of the characteristic polynomial provides the LRR coefficients, and vice versa.*

The following result [9] together with Remark 1 shows how to find the roots of the governing minimal LRR $s_{i+r} = \sum\limits_{k=1}^{r} a_k s_{i+r-k}$ using the common term of the time series S_N. Let $\mu_1, \dots, \mu_p$ be the roots of the LRR characteristic polynomial with multiplicities $k_1, \dots, k_p$. The time series S_N satisfies the LRR if and only if $s_n = \sum\limits_{m=1}^{p} \left(\sum\limits_{j=0}^{k_m - 1} c_{mj} n^j \right) \mu_m^n$, where $c_{mj} \in \mathbb{C}$ depends on $s_1, \dots, s_r$.

Example 1. *Consider the time series* $\mathsf{S}_N = (s_1, \dots, s_N)$ *with* $s_n = A\cos(2\pi\omega n + \phi)$, $\omega \in \left(0, \frac{1}{2}\right)$. *Since* $A\cos(2\pi\omega n + \phi) = Ae^{-\mathrm{i}\phi} e^{-\mathrm{i}\cdot 2\pi\omega n}/2 + Ae^{\mathrm{i}\phi} e^{\mathrm{i}\cdot 2\pi\omega n}/2$, *we have* $\mu_1 = e^{-\mathrm{i}\cdot 2\pi\omega}$, $\mu_2 = e^{\mathrm{i}\cdot 2\pi\omega}$. *Therefore, the characteristic polynomial is* $P_r(\mu) = (\mu - \mu_1)(\mu - \mu_2) = (\mu - e^{-\mathrm{i}\cdot 2\pi\omega n})(\mu - e^{\mathrm{i}\cdot 2\pi\omega n}) = \mu^2 - 2\mu\cos 2\pi\omega + 1$. *Thus,* $a_1 = 2\cos 2\pi\omega$, $a_2 = -1$ *and* $s_{i+2} = 2s_{i+1}\cos 2\pi\omega - s_i$.

Remark 2. *The method for constructing the min-norm LRR of a given order with the given signal roots is described in [10]. This method will be used in the algorithm proposed in Section 3.2.*

2.2. Harmonic Signal with Time-Varying Frequency: Instantaneous Frequency

In this paper, the basic form of signals will be the discrete-time version of

$$s(t) = \sum_{i=1}^{p} \cos(2\pi\omega_i(t) + \phi_i), \tag{2}$$

where $\omega_i(t)$, $i = 1, \dots, p$, are slowly changing functions. Note that if $\omega_i(t)$ are linear functions, the signal satisfies an LRR; see [4] (Section 2.2) and Remark 1.

Let $s(t) = \cos(2\pi\psi(t) + \phi)$. The instantaneous frequency of the signal $s(t)$ is defined as $\omega_{\text{ins}}(t) = \psi'(t)$. The instantaneous period is the function $T_{\text{ins}}(t) = \frac{1}{\omega_{\text{ins}}(t)}$. If $\omega_{\text{ins}}(a) = 0$ for some a, we put $T_{\text{ins}}(a) = +\infty$. The frequency range of the signal S_N is the range $\omega_{\text{ins}}([1, N])$, that is, the image of $[1, N]$.

Example 2. *For the signal* $s(t) = \cos\left(2\pi\left(\frac{t}{100}\right)^2\right)$*, the instantaneous frequency* $\omega_{\text{ins}}(t) = \frac{t}{5000}$ *is a linear function; the frequency range equals* $\omega_{\text{ins}}([1,\, N]) = \left[\frac{1}{5000}, \frac{N}{5000}\right]$*. For the signal* $s(t) = \cos\left(\frac{2\pi}{20}\left(t + 5\sin\frac{2\pi t}{100}\right)\right)$*, the instantaneous frequency equals* $\omega_{\text{ins}}(t) = \frac{1}{20}\left(1 + \frac{5\cdot 2\pi}{100}\cos\frac{2\pi t}{100}\right)$ *and is a periodic function with period equal to* $T = 100$*; the frequency range is*

$$\omega_{\text{ins}}([1, N]) \subset \left[\frac{1}{20}\left(1 - \frac{\pi}{10}\right), \frac{1}{20}\left(1 + \frac{\pi}{10}\right)\right].$$

We assume that on short segments of time, the signals considered in Example 2 are well-enough approximated by a sinusoid with a frequency equal to the instantaneous frequency in the middle of the segment. This assumption will be used for the construction of the signal-forecasting algorithm.

2.3. SSA, Signal Subspace and Recurrent SSA Forecasting

In the version for signal extraction, SSA has two parameters, the window length L, $1 < L < N$, where N is the time-series length, and the number of elementary components r. At the first step of SSA, the trajectory matrix of size $L \times K$, where $K = N - L + 1$, is constructed and then decomposed into elementary components using the SVD. The leading r SVD components are used for the estimation of the signal and the signal subspace basis, which is used for constructing the forecasting LRR. Here is the scheme of the method given in [6]:

$$\mathsf{X} \xrightarrow[\boxed{L}]{\mathcal{T}} \mathbf{X} = \begin{pmatrix} x_1 & x_2 & \dots & x_K \\ x_2 & x_3 & \dots & x_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_L & x_{L+1} & \dots & x_N \end{pmatrix} \xrightarrow[\boxed{r}]{\text{SVD:}(\sqrt{\lambda_m}, U_m, V_m),\ \Pi_r}$$

$$\begin{cases} \mathcal{L}_r = \operatorname{span}(U_1, \dots, U_r) \\ \text{is the signal space;} \\ \Pi_r \text{ is the projector on } \mathcal{L}_r; \\ \widehat{\mathbf{S}} = \sum_{m=1}^{r} U_m (\mathbf{X}^{\mathrm{T}} U_m)^{\mathrm{T}} = \Pi_r \mathbf{X}. \end{cases} \xrightarrow{\Pi_{\mathcal{H}}} \widetilde{\mathbf{S}} = \begin{pmatrix} \tilde{s}_1 & \tilde{s}_2 & \dots & \tilde{s}_K \\ \tilde{s}_2 & \tilde{s}_3 & \dots & \tilde{s}_{K+1} \\ \vdots & \vdots & \ddots & \vdots \\ \tilde{s}_L & \tilde{s}_{L+1} & \dots & \tilde{s}_N \end{pmatrix} \xrightarrow{\mathcal{T}^{-1}} \widetilde{\mathsf{S}}.$$

Thus, a concise form of the SSA algorithm for signal extraction is

$$\widetilde{\mathsf{S}} = \mathcal{T}^{-1} \circ \Pi_{\mathcal{H}} \circ \Pi_r \circ \mathcal{T}(\mathsf{X}),$$

where $\Pi_{\mathcal{H}}$ is the projector to the set of Hankel matrices, that is, the set of trajectory matrices.

The L-rank of a time series is the rank of its trajectory matrix $\mathcal{T}(\mathsf{X})$, or, equivalently, the dimension of the column space of $\mathcal{T}(\mathsf{X})$. For infinite time series and $L > r$, the rank is equal to the order of the minimal LRR governing the time series. For example, the rank of the signal $\mathsf{S}_\infty = (s_1, s_2, \dots)$ with the common term $s_n = A\cos(2\pi\omega n + \phi)$ and $\omega \in \left(0, \frac{1}{2}\right)$ equals two.

The construction of the forecasting min-norm LRR of order $L - 1$ based on the SSA decomposition follows the formula for the LRR coefficients $\mathcal{R}$ [4] (Equation (2.1)):

$$\mathcal{R} = \frac{1}{1 - \nu^2} \sum_{i=1}^{r} \pi_i \underline{U_i} =: (a_{L-1}, \dots, a_1), \tag{3}$$

where $\underline{U_i} \in \mathbb{R}^{L-1}$ is the vector U_i without the last coordinate, which is denoted by π_i. The forecast is constructed as $\tilde{s}_{N+1} = \sum_{i=1}^{L-1} a_i \tilde{s}_{N+1-i}$.

In addition to forecasting, the SSA decomposition allows one to estimate the signal roots. Let us describe the ESPRIT (see [7] (Algorithm 3.3) and [11]) for signal-root estimation. We define $\mathbf{U} := [U_1 : \dots : U_r]$, $\overline{\mathbf{U}}$ the matrix $\mathbf{U}$ without the first row and $\underline{\mathbf{U}}$ the matrix $\mathbf{U}$ without the last row. The ESPRIT estimates of the signal roots $\{\mu_k\}_{k=1}^{r}$ are the eigenvalues of a matrix that is an approximate solution of the equation $\underline{\mathbf{U}}\mathbf{D} = \overline{\mathbf{U}}$; e.g., $\mathbf{D} = (\underline{\mathbf{U}}^{\mathrm{T}}\underline{\mathbf{U}})^{-1}\underline{\mathbf{U}}^{\mathrm{T}}\overline{\mathbf{U}}$ for the LS-ESPRIT version.

3. Signal Forecasting by Forecasting of Local LRRs

3.1. General Model of Signals

Let us describe a general model of time series, for which the developed approach to forecasting will be applied. Consider the signal S_N and take some natural Z, $1 < Z < N$. For a time series Y, we denote $\mathsf{Y}_{A,B}$ the series $(y_A, \dots, y_B)$.

The signal model is

$$s_n = \sum_{j=1}^{p} \rho_j^n(n) \cos(2\pi\psi_j(n) + \varphi_j), \tag{4}$$

where we assume that

1. For the time series S, on its sequential segments $\mathsf{S}_{i,i+Z-1}$ of length Z, $i = 1, \dots, N - Z + 1$, every summand in (4) is well-approximated by a series in the form $s_n^{\text{approx}} = \rho_j^n(n_0) \cos(2\pi\omega_{j,0} \cdot n + \phi)$, where $\omega_{j,0} = \psi_j'(n_0)$ and $n_0 = n_0(i)$ is the middle point of the segment.
2. The series $\rho_j(n)$ and $\psi_j'(n)$ behave regularly in n, $0 < \psi_j'(n) < 0.5$ and there exist methods that can forecast such kinds of series.

To construct a forecast s_{N+1} one needs to find the instantaneous frequencies at the point $N + 1$.

3.2. Algorithm LocLRR SSA Forecast

Hereinafter, we will consider the model $\mathsf{X}_N = \mathsf{S}_N + \mathsf{R}_N$, where $\mathsf{S}_N = (s_1, \dots, s_N)$ satisfies the conditions described above and R_N is white Gaussian noise with zero mean and standard deviation σ. Let $\mathsf{Z}_{(i)} = \mathsf{X}_{i,i+Z-1}$, $i = 1, \dots, W$, where $W = N - Z + 1$. The estimates of instantaneous roots, frequencies, moduli and LRR coefficients will be enumerated according to the middle of the local segment. In particular, denote $\{a_k^{(i)}\}_{k=1}^r$ the coefficients of the minimal LRR that approximates the local segment of the series $S_{i-[Z/2],i+[Z/2]-1}$ with the center in s_i.

The local segment of length Z has the structure depicted in Figure 1, where the middle of $\mathsf{Z}_{(i_0)}$ is $N + 1$.

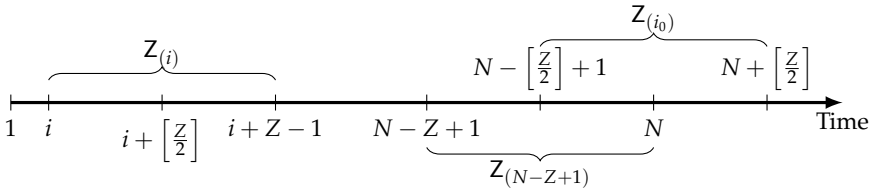


Figure 1. Scheme of moving segments.

3.2.1. Scheme

In the scheme below, we consider $j = 1, \dots, p'$, where p' is the number of signal roots with nonnegative imaginary parts (note that signal roots with negative imaginary parts are conjugate to signal roots with positive imaginary parts):

$$\begin{array}{ccccccc} \mathsf{X}_N & & \{\rho_j^{(i+[Z/2])}\}_{i=1}^W & \xrightarrow{\text{FOR MODs}} & \{\widetilde{\rho}_j^{(N+l)}\}_{l=1}^M & & \\ \downarrow \mathsf{Z} & & \uparrow & & \downarrow & & \\ \{\mathsf{Z}_{(i)}\}_{i=1}^W & \xrightarrow[L,r]{ESPRIT} & \{\mu_j^{(i+[Z/2])}\}_{i=1}^W & & \{\widetilde{\mu}_j^{(N+l)}\}_{l=1}^M & \to \{\widetilde{\mathcal{R}}_{N+l}\}_{l=1}^M. & \\ & & \downarrow & & \uparrow & & \\ & & \{\omega_j^{(i+[Z/2])}\}_{i=1}^W & \xrightarrow{\text{FOR ARGs}} & \{\widetilde{\omega}_j^{(N+l)}\}_{l=1}^M & & \end{array}$$

Here $\{\widetilde{\mathcal{R}}_{N+l}\}_{l=1}^M$, where $\widetilde{\mathcal{R}}_k = (\tilde{a}_r, \dots, \tilde{a}_1)$ is the sequence of coefficients of the forecasting minimal LRRs. If the required length m of the forecasting LRRs is larger than r, then we lengthen each $\widetilde{\mathcal{R}}_{N+l}$, $l = 1, \dots, M$, to $\widetilde{\mathcal{R}}_{N+l}^{(m)}$ of the min-norm LRR of order m; see Remark 2.

The result is a sequence of coefficients of the forecasting min-norm LRRs $\{\widetilde{\mathcal{R}}^{(m)}_{N+l}\}_{l=1}^{M}$ of order m. To obtain the results, the series of moduli and frequencies should be forecasted using some algorithms FOR MODs and FOR ARGs for each j.

After the sequence of coefficients is constructed, the forecasting values $\widetilde{\mathsf{S}}_{N+1,N+M} = (y_{N+1}, \ldots, y_{N+M})$ are taken from the time series

$$y_n = \begin{cases} \tilde{s}_n, & n = 1, \ldots, N, \\ \sum\limits_{k=1}^{m} \widetilde{a}_k^{(n)} y_{n-k}, & n = N+1, \ldots, N+M, \end{cases}$$

where $\tilde{s}_n$, $n = 1, \ldots, N$, is the signal estimate obtained using SSA.

3.2.2. Algorithm in Detail

Let us formally describe the algorithm for forecasting the minimal LRRs (Algorithm 1).

Algorithm 1: LocLRR SSA Forecast

Input:

- time series $\mathsf{X}_N = (x_1, \ldots, x_N)$,
- forecast length M,
- local segment length Z,
- window length L,
- number r of leading eigentriples of the trajectory matrices of the local segments that are used to find the estimates of the signal roots,
- length m of the forecasting LRRs,
- algorithm for forecasting the root moduli FOR MODs,
- algorithm for forecasting the instantaneous frequencies FOR ARGs.

Steps:

1. For each segment $\mathsf{Z}_{(i)}$, $i = 1, \ldots, W$, where $W = N - Z + 1$, estimate the signal roots $\{\mu_j^{(i+[Z/2])}\}_{j=1}^{r}$ of the approximating series using ESPRIT with window length L and signal rank r. Suppose that the estimates of the roots form complex-conjugate pairs. Choose the roots with positive argument $\{\mu_j^{(i+[Z/2])}\}_{j=1}^{p}$, $i = 1, \ldots, W$, $p = \frac{r}{2}$.
2. Arrange the first set of roots $\{\mu_j^{(1+[Z/2])}\}_{j=1}^{p}$ in descending argument order.
3. Order the sets of roots $\{\mu_j^{(i+[Z/2])}\}_{j=1}^{p}$, $i = 2, \ldots, W$, so that the sum
$$\sum_{j=1}^{p} \left| \mu_j^{(i+[Z/2])} - \mu_j^{(i+[Z/2]-1)} \right|$$
is minimal among all possible permutations of $\{\mu_j^{(i+[Z/2])}\}_{j=1}^{p}$.
4. For each $i = 1, \ldots, W$ and $j = 1, \ldots, p$, calculate the moduli $\rho_j^{(i+[Z/2])} = |\mu_j^{(i+[Z/2])}|$ and the frequencies $\omega_j^{(i+[Z/2])} = \frac{\operatorname{Arg} \mu_j^{(i+[Z/2])}}{2\pi}$. For $j = 1, \ldots, p$, set the series $P_{(j)} = \left(\rho_j^{(1+[Z/2])}, \ldots, \rho_j^{(N-[Z/2]+1)}\right)$ and $\Omega_{(j)} = \left(\omega_j^{(1+[Z/2])}, \ldots, \omega_j^{(N-[Z/2]+1)}\right)$.
5. Using the algorithms FOR MODs and FOR ARGs, for each $j = 1, \ldots, p$, construct the forecast $\left(\widetilde{\rho}_j^{(N+1)}, \ldots, \widetilde{\rho}_j^{(N+M)}\right)$ of the series $P_{(j)}$ and the forecast $\left(\widetilde{\omega}_j^{(N+1)}, \ldots, \widetilde{\omega}_j^{(N+M)}\right)$ of $\Omega_{(j)}$, respectively.
6. Using the obtained forecasts of frequencies and moduli, calculate the roots $\widetilde{\mu}_j^{(n)} = \widetilde{\rho}_j^{(n)} \exp(\mathrm{i}\, 2\pi \widetilde{\omega}_j^{(n)})$, $n = N+1, \ldots, N+M$, $j = 1, \ldots, p$, supplement them by their complex conjugates and then, using the relation between characteristic polynomials and LRRs (see Remark 1), find the sequence $\{\widetilde{\mathcal{R}}_{N+j}\}_{j=1}^{M}$ of the LRR coefficients, $\widetilde{\mathcal{R}}_{N+j} = \left(\widetilde{a}_r^{(N+j)}, \ldots, \widetilde{a}_1^{(N+j)}\right)$.

Output: The sequence $\{\widetilde{\mathcal{R}}_{N+l}\}_{l=1}^{M}$ of coefficients of minimal LRRs of order r approximately governing the future signal segments $\mathsf{S}_{l+N-[Z/2]-1, l+N+[Z/2]}$.

Remark 3. *If real-valued roots are obtained on some segments of* $\mathsf{Z}_{(i)}$, $i = 1, \dots, W$, *we replace the values of the roots with missing values. That is, for* $i = 1, \dots, W$ *and* $j = 1, \dots, r$ *such that* $\mu_j^{(i)} \in \mathbb{R}$ *we put* $\mu_j^{(i)} :=$ NA (*not available*). *Possible gaps in the series of frequency estimates can be filled in; e.g., one can fill them with the iterative gap-filling method [12]; see also the description of the igapfill algorithm in [7] (Algorithm 3.7).*

An appropriate choice of the algorithms FOR MODs and FOR ARGs depends on the form of the frequency modulation.

Finally, each LRR of $\{\widetilde{\mathcal{R}}_{N+j}\}_{j=1}^{M}$ is enlarged to a min-norm LRR of length m with coefficients $\{\widetilde{\mathcal{R}}_{N+l}^{(m)}\}_{l=1}^{M}$ (see Remark 2) and this enlarged LRR is used for the prediction of s_{N+l}.

4. Examples

4.1. Description

In this section, we demonstrate the forecasting using the LocLRR SSA Forecast algorithm. The following types of time series were considered.

4.1.1. Sinusoid with Linearly Modulated Frequency

The signal has the form

$$s(t) = \cos(2\pi(at)^2), \quad a \neq 0.$$

The instantaneous frequency is $\omega_{\text{ins}}(t) = 2a^2t$; the frequency range is $\omega_{\text{ins}}([1, N]) = [2a^2, 2a^2N]$. We will consider such values of the parameter a and series length N at which the frequency range $\omega_{\text{ins}}([1, N]) \subset (0, 0.5)$. Since the instantaneous frequency is a linear function, we will take the linear-regression-prediction algorithm as FOR ARGs.

4.1.2. Sinusoid with Sinusoidal Frequency

The signal is

$$s(t) = \cos(2\pi\omega_{\text{ext}}(t + b\sin 2\pi\omega_{\text{int}}t)), \quad b > 0,$$

$\omega_{\text{ext}}, \omega_{\text{int}} \in (0, 0.5)$, where ω_{int} is much smaller than ω_{ext}. The instantaneous frequency equals $\omega_{\text{ins}}(t) = \omega_{\text{ext}}(1 + 2\pi\omega_{\text{int}}b\cos 2\pi\omega_{\text{int}}t)$; the frequency range is $\omega_{\text{ins}}([1, N]) = [\omega_{\text{ext}}(1 - 2\pi\omega_{\text{int}}b), \omega_{\text{ext}}(1 + 2\pi\omega_{\text{int}}b)]$. We will consider values of signal parameters and series lengths such that $\omega_{\text{ins}}([1, N]) \subset (0, 0.5)$. The rank of the time series of the series Ω_N of instantaneous frequencies with terms $\omega_n = \omega_{\text{ins}}(n)$, $n = 1, \dots, N$, is equal to 3. Therefore, we take the recurrent SSA forecasting algorithm with $r = 3$ as FOR ARGs; the window length L is chosen to be half of the length of the frequency-estimation series following the general recommendations.

Since we do not consider time-varying amplitude modulation in the examples, we take the forecast of moduli series using the average value over local intervals as algorithm FOR MODs.

4.1.3. Sum of Sinusoids

The signal is

$$s(t) = \sum_{j=1}^{p_1} \cos(2\pi(a_jt)^2) + \sum_{k=1}^{p_2} \cos\left(2\pi\omega_k^{\text{ext}}\left(t + b_k\sin 2\pi\omega_k^{\text{int}}t\right)\right), \tag{5}$$

$a_j \neq 0$, $j = 1, \dots, p_1$, $\omega_k^{\text{ext}}, \omega_k^{\text{int}} \in (0, 0.5)$, $b_k > 0$, $k = 1, \dots, p_2$. We will consider the signal parameters and the time-series length N such that the frequency ranges of the summands are not mutually intersected and the frequency range of each summand belongs to the interval $(0, 0.5)$.

The frequencies corresponding to the signal summands of $s(t)$ will be forecasted either with linear regression or with SSA, based on the type of obtained estimates of the instantaneous frequencies on the local segments. The moduli will be predicted using the average of the estimates over the local segments.

In numerical examples, we will consider the signal parameters such that the instantaneous frequencies are slow-varying functions. Since the examples under consideration satisfy the general time-series model (Section 3.1), the LocLRR SSA-forecast algorithm can be used for forecasting.

In real-life problems, the value of the optimal LRR order m can be chosen based on the training data. In the model examples, the value of the optimal LRR order m will be chosen by trying all possible values of m in the range from r to Z and comparing the mean squared errors (MSE) of the predictions.

The following approach was used for choosing the length Z of local segments. We take such a value of the parameter Z that most of the local segments $\mathsf{Z}_{(i)}$, $i = 1, \ldots, W$, contain at least 2–3 instantaneous periods of each summand of the signal satisfying model (5). For small Z, we obtain estimates of the instantaneous frequency with large variability, whereas for large Z, we have a considerable bias. The necessary condition for appropriate values of Z is that there are no (or a few) segments providing real-valued roots.

4.2. Numerical Experiments

Consider the following numerical examples:

- Sinusoid with linearly modulated frequency,

$$s_n = \cos\left(2\pi\left(\frac{n}{100}\right)^2\right)$$

(denoted by $\cos(\mathsf{n}^2)$);

- Sinusoid with sinusoidal frequency modulation,

$$s_n = \cos\left(\frac{2\pi}{20}\left(n + 5\sin\frac{2\pi n}{100}\right)\right)$$

(denoted by $\cos\sin(\mathsf{n})$);

- Sum of sinusoids with linear and sinusoidal frequency modulations,

$$s_n = \cos\left(2\pi\left(\frac{n}{100}\right)^2\right) + \cos\left(\frac{2\pi}{10}\left(n + \sin\frac{2\pi n}{100}\right)\right) \tag{6}$$

(denoted by $\cos(\mathsf{n}^2) + \cos\sin(\mathsf{n})$);

- Sum of two sinusoids with sinusoidal frequency modulation,

$$s_n = \cos\left(\frac{2\pi}{20}\left(n + \sin\frac{2\pi n}{100}\right)\right) + \cos\left(\frac{2\pi}{10}\left(n + 2\sin\frac{2\pi n}{140}\right)\right)$$

(denoted by $\cos\sin(\mathsf{n}) + \cos\sin(\mathsf{n})$).

We compare the proposed LocLRR SSA-forecast algorithm (denoted by "alg") with two simple methods:

- Forecasting by constant, which forecasts by zero, since we consider time series with zero average (denoted by 'by 0').
- Forecasting using the last local segment, which is performed with the min-norm LRR computed using the roots of the last local segment $\mathsf{Z}_{(N-Z+1)} = (x_{N-Z+1}, \ldots, x_N)$ (denoted by 'last').

Let us consider accuracies of $M = 30$ step ahead forecasts for time series of length $N = 300$. For each example, the prediction is performed for the pure signal and the noisy signal, where the noise is white Gaussian with standard deviation $\sigma = 0.25$. In the noisy case, a sample of size $P = 100$ is used for the estimation of accuracy.

Let $\{\mathsf{X}_N^{(i)}\}_{i=1}^P$ be the time-series sample. The prediction error is estimated as RMSE $= \sqrt{\frac{1}{P}\sum_{i=1}^{P} \mathrm{MSE}(\widetilde{\mathsf{S}}_{N+1,N+M}^{(i)}, \mathsf{S}_{N+1,N+M})}$.

The results are shown in Table 1, where the best results are highlighted in bold. They confirm the advantage of the proposed method over the simple methods under consideration.

Table 1. RMSE of forecasts; 'alg' is the proposed algorithm; m is the optimal length of the forecasting LRRs.

Signal $\mathbf{S}_N$	σ	by 0	Last		Alg	
			RMSE	m	RMSE	m
$\cos(\mathsf{n}^2)$	0	0.689	0.717	2	**0.014**	3
	0.25	0.733	0.754	5	**0.135**	11
$\cos\sin(\mathsf{n})$	0	0.698	0.309	2	**0.097**	2
	0.25	0.741	0.438	5	**0.232**	6
$\cos(\mathsf{n}^2) + \cos\sin(\mathsf{n})$	0	1.060	0.880	6	**0.184**	4
	0.25	1.089	0.958	10	**0.295**	12
$\cos\sin(\mathsf{n}) + \cos\sin(\mathsf{n})$	0	0.873	0.587	4	**0.191**	5
	0.25	0.908	0.656	28	**0.291**	15

4.3. Detailed Example

To demonstrate the approach more clearly, let us consider the example (6) without noise; see Figure 2. Take $Z = 61$, $L = 30$, $r = 4$.

The results of forecasting the signal root corresponding to the first summand are shown in Figure 3, and the results of forecasting the signal root corresponding to the second summand can be seen in Figure 4. The frequency ranges of modulations in the summands are approximately $[0, 0.06]$ and $[0.094, 0.106]$, respectively. The forecasts are depicted in Figure 5 (forecasting with the last segment) and Figure 6 (forecasting with the proposed algorithm). Since there is no noise, the optimal LRR length m is small; here, $m = 4$.

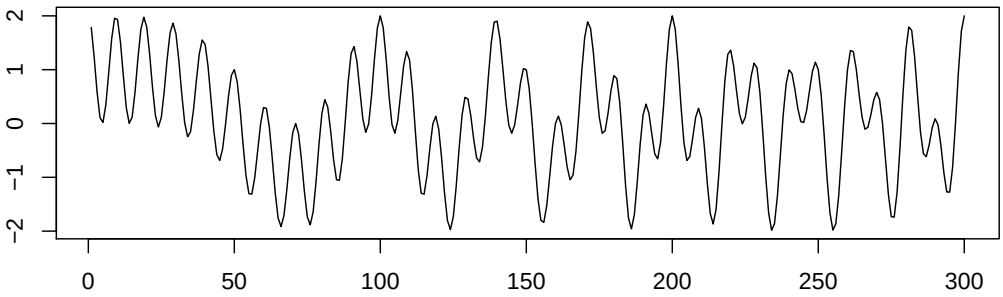


Figure 2. Initial signal $s_n = \cos\left(2\pi\left(\frac{n}{100}\right)^2\right) + \cos\left(\frac{2\pi}{10}\left(n + \sin\frac{2\pi n}{100}\right)\right)$.

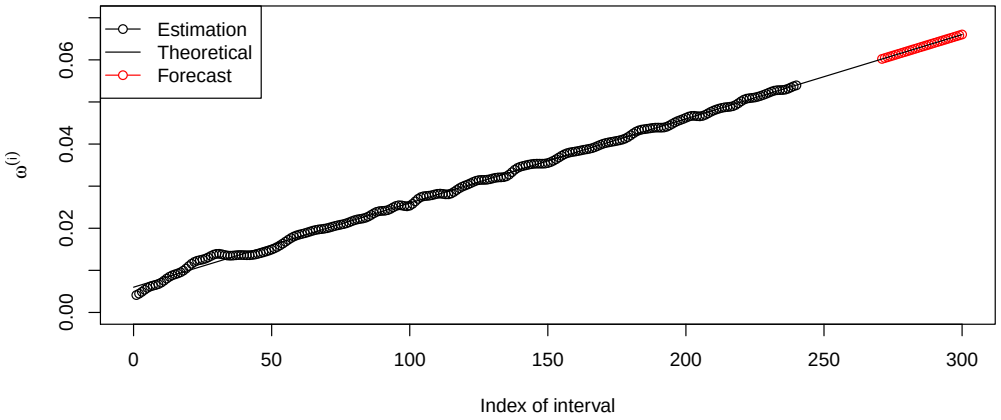


Figure 3. Forecasting the series of linear instantaneous frequencies for the summand $\cos\left(2\pi\left(\frac{n}{100}\right)^2\right)$.

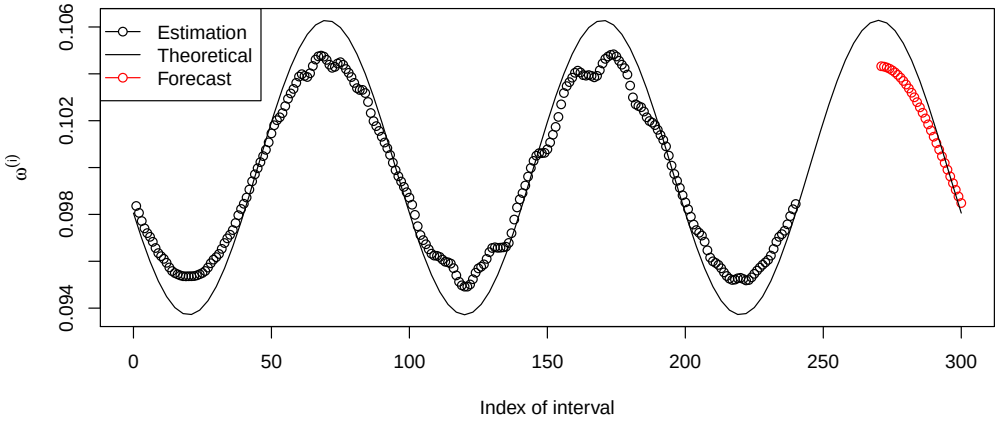


Figure 4. Forecasting the series of sinusoidal instantaneous frequencies for the summand $\cos\left(\frac{2\pi}{10}\left(n+\sin\frac{2\pi n}{100}\right)\right)$.

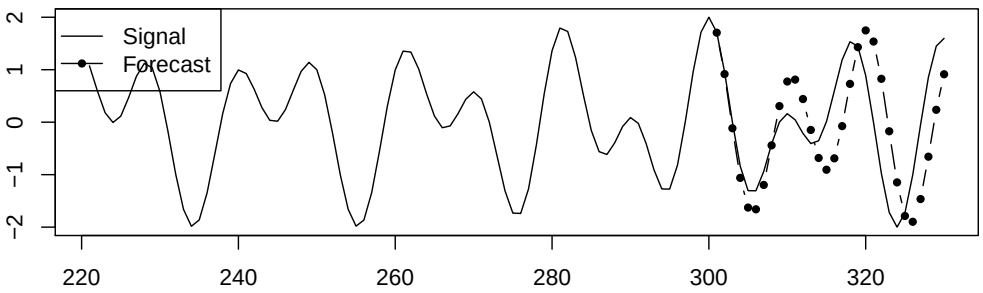


Figure 5. Forecasting using the last local segment, RMSE $= 0.88$. A shift is clearly seen.

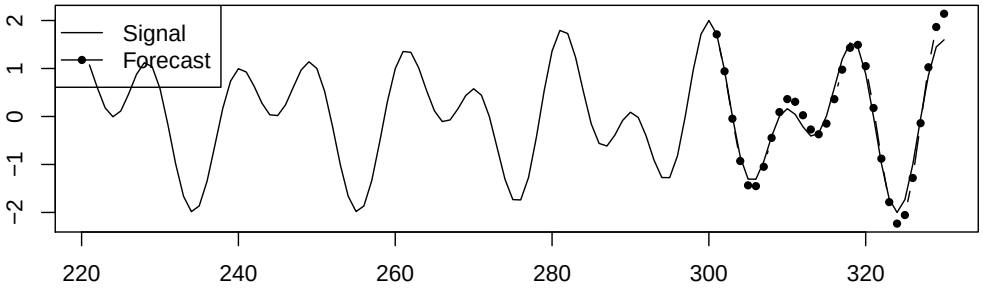


Figure 6. Forecasting using LocLRR SSA forecast, RMSE $= 0.2$. There is no shift.

5. Conclusions

In this paper, we proposed a method for forecasting time series that extends the capabilities of SSA and allows one to predict time series in which the signal just locally satisfies LRRs. A regular behaviour of the coefficients of the governed LRRs was assumed. In [6] (page 9), it was stated for the considered type of time series that "[t]he problem is how to forecast the extracted signal, since its local estimates may have different structures on different time intervals. Indeed, by using local versions of SSA, we do not obtain a common nonlinear model but instead we have a set of local linear models". In this paper, we proposed an answer to the problem of prediction of local structures of time series for some class of signals.

We constructed an algorithm for predicting local structures of time series that are the sum of frequency-modulated sinusoids and showed that the proposed forecasting method gives reasonable results for the cases of linear and sinusoidal frequency modulations.

Certainly, the considered comparison with a couple of simple methods is not enough; a more extensive comparison should be performed in the future. However, the results of this work show that the proposed approach based on the prediction of the coefficients of LRRs is promising.

Author Contributions: Conceptualization, N.G.; methodology, N.G. and E.S.; software, E.S.; validation, N.G. and E.S.; formal analysis, N.G. and E.S.; investigation, N.G. and E.S.; resources, N.G.; data curation, N.G. and E.S.; writing—original draft preparation, N.G. and E.S.; writing—review and editing, N.G. and E.S.; visualization, N.G. and E.S.; supervision, N.G.; project administration, N.G.; funding acquisition, N.G. All authors have read and agreed to the published version of the manuscript.

Funding: The work is supported by the Russian Science Foundation (project No. 23-21-00222).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study are openly available. The R-code for data generation and result replication can be found at https://zenodo.org/record/8087608, doi 10.5281/zenodo.8087608.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Broomhead, D.; King, G. Extracting qualitative dynamics from experimental data. *Physica D* **1986**, *20*, 217–236. [CrossRef]
2. Vautard, R.; Ghil, M. Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Physica D* **1989**, *35*, 395–424. [CrossRef]
3. Elsner, J.B.; Tsonis, A.A. *Singular Spectrum Analysis: A New Tool in Time Series Analysis*; Plenum Press: New York, NY, USA, 1996.
4. Golyandina, N.; Nekrutkin, V.; Zhigljavsky, A. *Analysis of Time Series Structure: SSA and Related Techniques*; Chapman & Hall/CRC: Boca Raton, FL, USA, 2001.
5. Golyandina, N.; Zhigljavsky, A. *Singular Spectrum Analysis for Time Series*, 2nd ed.; Springer Briefs in Statistics; Springer: Berlin/Heidelberg, Germany, 2020.
6. Golyandina, N. Particularities and commonalities of singular spectrum analysis as a method of time series analysis and signal processing. *Wiley Interdiscip. Rev. Comput. Stat.* **2020**, *12*, e1487. [CrossRef]
7. Golyandina, N.; Korobeynikov, A.; Zhigljavsky, A. *Singular Spectrum Analysis with R*; Springer: Berlin/Heidelberg, Germany, 2018.
8. Leles, M.; Sansão, J.; Mozelli, L.; Guimarães, H. Improving reconstruction of time-series based in Singular Spectrum Analysis: A segmentation approach. *Digit. Signal Process.* **2018**, *77*, 63–76. [CrossRef]
9. Hall, M.J. *Combinatorial Theory*; Wiley: New York, NY, USA, 1998.
10. Usevich, K. On signal and extraneous roots in Singular Spectrum Analysis. *Stat. Interface* **2010**, *3*, 281–295. [CrossRef]
11. Roy, R.; Kailath, T. ESPRIT-estimation of signal parameters via rotational invariance techniques. *IEEE Trans. Acoust. Speech Signal Process.* **1989**, *37*, 984–995. [CrossRef]
12. Kondrashov, D.; Ghil, M. Spatio-temporal filling of missing points in geophysical data sets. *Nonlinear Process. Geophys.* **2006**, *13*, 151–159. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.