ISSN 1063-4541, Vestnik St. Petersburg University, Mathematics, 2023, Vol. 56, No. 1, pp. 35–45. © Pleiades Publishing, Ltd., 2023.
Russian Text © The Author(s), 2023, published in Vestnik Sankt-Peterburgskogo Universiteta: Matematika, Mekhanika, Astronomiya, 2023, Vol. 10, No. 1, pp. 48–61.

MATHEMATICS

Remark on the Accuracy of Recurrent Forecasting in Singular Spectrum Analysis

V. V. Nekrutkin

St. Petersburg State University, St. Petersburg, 199034 Russia
e-mail: vnekr@statmod.ru

Received July 18, 2022; revised August 29, 2022; accepted September 8, 2022

Abstract—Within the series of singular spectrum analysis (SSA) methods, there exist several versions of forecasting algorithms for signals corrupted by additive noise. In this paper, a technique is proposed to estimate the asymptotic accuracy of the recurrent version of such forecasting when the length of a series tends to infinity. Most elements of this construction can be reduced to already studied and published results, although some of them are hard to implement in specific situations. The article brings together all these elements and augments and comments on them. Several examples of determining estimates of accuracy for a recurrent forecast are given for specific signals and noises. The computational experiments carried out confirm the theoretical results.

Keywords: signal processing, singular spectrum analysis, recurrent forecast, asymptotic analysis

DOI: 10.1134/S1063454123010090

1. INTRODUCTION

Several variants of application of singular spectrum analysis (SSA) to forecast signals corrupted by additive noise were considered in ([1], Ch. 2) (see also [2]). Among them, the main one is the so-called recurrent forecast. However, by all accounts, there are still no theoretical results concerning the forecast accuracy for large noise (not in the linearized formulation of the problem). In the present study, we make an attempt to fill this gap.

Without going into details (a full description and discussion can be found in Section 3 of this paper), we note that a recurrent forecast is carried out by a specially constructed linear recurrent formula applied to a signal reconstructed by the SSA method. In this sense, this method is similar to a usual linear forecast (see, for example, [3]). The differences, however, are significant, and the most important of them is that a linear forecast is applied to the entire observed series, whereas a recurrent forecast is applied to its "noisy" additive component (signal). That is why it is required to additionally estimate ("reconstruct") this signal, which is done by the SSA method. In addition, the coefficients of the recurrent formula in the forecasting under consideration are expressed in terms of the projection onto the signal subspace (disturbed by noise), which makes it possible to apply the theory of [4] to their estimation.

In this paper, we propose a method for estimating the accuracy of a recurrent forecast when the length of the series tends to infinity. We note that almost all components of this method are already known and published, although some of them are difficult to apply in specific situations. Let us describe these components.

First, these are the *verticality coefficients* introduced in ([1], § 5.2); they are described and illustrated in detail in Section 2 of this paper. The examples given show that the study of the asymptotic behavior of the verticality coefficients as the series length tends to infinity is quite realistic.

The second component is the *sine of the largest principal angle between perturbed and unperturbed linear subspaces of the signal* (see [4] for a detailed description and the corresponding mathematical technique). Numerous examples in [4] convince that the asymptotic behavior of this characteristic can also be studied for a wide class of signals and noise.

The last component is the *accuracy of signal reconstruction* by SSA. Although a general approach to estimating this accuracy was published in [4], the procedure is generally quite laborious, and at present there are only a few examples of its application.

In this paper, we bring together all these components and supplement and comment on them.

Section 2 of the paper is devoted to the definition and asymptotic properties of the verticality coefficients of signals governed by linear recurrent formulas (LRFs). In Section 3 we discuss the definition and properties of a recurrent forecast, as well as a method for estimating its accuracy.

In Subsection 3.4 we give examples of the application of this estimate, and, in the Appendix, the results of computational experiments illustrating the theoretical constructions.

The content and style of these examples are determined by the variants of signals and noise studied in previous works. The signals include series that have a simple structure from the viewpoint of SSA but, at the same time, are often encountered in practice as models of trends and periodic components.

We note that the approach to recurrent forecasting in this paper is somewhat different from that proposed in ([1], Ch. 2). Namely, here the so-called approximate LRF (see Subsection 3.1) is constructed using a window of length M, which may differ from the window length L used to obtain reconstruction errors. A similar approach was used in computational experiments in [5]. At the same time, in ([1], Ch. 2), the authors considered only the case of $L = M$, which is more convenient for practical implementation.

However, for $L \sim \alpha N$ and $M \sim \beta N$ with $\alpha, \beta \in (0, 1)$, the asymptotic result for forecast errors at the series length $N \to \infty$ will be largely similar.

2. VERTICALITY COEFFICIENTS AND LRF

2.1. Definition and Examples

Let $\mathfrak{U}$ be a linear subspace of $\mathbb{R}^M$ of dimension $d = \dim\mathfrak{U}$. We denote by Π the orthogonal projector onto $\mathfrak{U}$ and set $\mathfrak{e}_M = (0, 0, ..., 0, 1)^{\mathrm{T}} \in \mathbb{R}^M$.

We call the number $\vartheta = \|\Pi\mathfrak{e}_M\|$ the *verticality coefficient* of $\mathfrak{U}$ and note that ϑ is the cosine of the angle between the vector $\mathfrak{e}_M$ and the linear space $\mathfrak{U}$. If $\vartheta = 1$ (in other words, if $\mathfrak{e}_M \in \mathfrak{U}$), then $\mathfrak{U}$ is said to be *vertical*.

Let $P_1, ..., P_d$ be an orthonormal basis in $\mathfrak{U}$; then $\pi_i = (P_i, \mathfrak{e}_M)$ is the last coordinate P_i and

$$\Pi\mathfrak{e}_M = \sum_{i=1}^{d}\pi_i P_i \quad \text{and} \quad (\Pi\mathfrak{e}_M, \mathfrak{e}_M) = \sum_{i=1}^{d}\pi_i^2 = \vartheta^2. \tag{1}$$

The series $\{y_n, n \ge 1\}$ is said to be governed by a LRF with the coefficient vector $R = (c_{K-1}, ..., c_1)^{\mathrm{T}}$ if the following relation holds for $n \ge K$:

$$y_n = \sum_{k=1}^{K-1} c_k y_{n-k}.$$

Lemma 1. *We suppose that the series $F = (f_1, ..., f_n, ...)$ is governed by the LRF*:

$$f_n = \sum_{k=1}^{M-1} b_k f_{n-k}, \quad n \ge M. \tag{2}$$

For $1 \le \ell \le k$, we set $F_{\ell,k} = (f_\ell, ..., f_k)^{\mathrm{T}}$ and denote by $\mathrm{U}_0^{\perp}(M)$ a linear space with the generators $\{F_{n,n+M-1}, n \ge 1\}$. Then $\mathrm{U}_0^{\perp}(M)$ is not vertical.

Proof. We set $R(M) = (b_{M-1}, ..., b_1)^{\mathrm{T}}$. Then (2) takes the form

$$(Z_M, F_{n-M+1,n}) = 0 \quad \text{with} \quad Z_M = \begin{pmatrix} -R(M) \\ 1 \end{pmatrix}, \quad n \ge M;$$

i.e., Z_M is orthogonal to $\mathrm{U}_0^{\perp}(M)$. Since $(Z_M, \mathfrak{e}_M) = 1$, it follows that $\mathfrak{e}_M \notin \mathrm{U}_0^{\perp}(M)$.

Now, we suppose that the series $F = (f_1, ..., f_n, ...)$ is governed by a minimal LRF of order d,

$$f_n = \sum_{k=1}^{d} a_k f_{n-k}, \quad n > d, \tag{3}$$

and the term *minimal* means that there does not exist a LRF of order $d' < d$ that governs the series F.

Taking $M \geq d$, as before, we denote by $\mathrm{U}_0^{\perp}(M)$ the linear vector space spanned by $\{F_{n,n+M-1},\ n \geq 1\}$. Moreover, let $\mathbf{P}_0^{\perp}(M)$ be the orthogonal projector $\mathbb{R}^M \mapsto \mathrm{U}_0^{\perp}(M)$. Then $\dim \mathrm{U}_0^{\perp}(M) = d$ and $\vartheta_M = \left\|\mathbf{P}_0^{\perp}(M) e_M\right\| < 1$.

We consider $P_1, \ldots, P_d$, which is a basis in $\mathrm{U}_0^{\perp}(M)$, and denote $\mathbf{X} = [P_1 : \ldots : P_d]$. Then it is easy to see that

$$\vartheta^2(M) = e_M^{\mathrm{T}} \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}} e_M. \tag{4}$$

Using (4), we present examples of the asymptotic behavior of ϑ_M^2 as $M \to \infty$ for several variants of f_n, namely, for exponential, polynomial, and trigonometric series. In this case, we will omit some elementary but cumbersome calculations.

2.1.1. Verticality coefficients. Examples

1. *Exponential series.* Here

$$f_n = \sum_{i=1}^{p} \beta_i a_i^n \tag{5}$$

with $a_i > 1$ and $\beta_i \neq 0$. Thus, $\dim \mathrm{U}_0^{\perp}(M) = p$ for $M \geq p$, and the vectors $P_\ell = A_{M,\ell}/\|A_{M,\ell}\|$ with $A_{M,\ell} = (1, a_\ell, \ldots, a_\ell^{M-1})^{\mathrm{T}}$ and $\ell = 1, \ldots, p$ form the basis of the space $\mathrm{U}_0^{\perp}(M)$.

We denote $\mathbf{X} = [P_1 : \ldots : P_p]$. If $M \to \infty$, then $\mathbf{X}^{\mathrm{T}}\mathbf{X} \to \mathbf{C}$, where $\mathbf{C}$ is a $p \times p$ matrix with entries $c_{ij} = \sqrt{(a_i^2-1)(a_j^2-1)/(a_i a_j - 1)}$. Similarly, $e_M^{\mathrm{T}}\mathbf{X} \to \left(\sqrt{1-a_1^{-2}}, \ldots, \sqrt{1-a_p^{-2}}\right) := D^{\mathrm{T}}$ and $\vartheta_M^2 \to \vartheta_\infty^2 = D^{\mathrm{T}}\mathbf{C}^{-1}D$. When $p = 1$, the equality $\vartheta_\infty^2 = 1 - 1/a_1^2$ holds.

2. *Scheme of series for exponential series.* This case was studied in detail from the viewpoint of SSA in [6].

For $T > 0$, we consider the scheme of series for the series $f_n = f_n^{(N)} = a^{nT/N}$, where $a > 1$, $n = 1, \ldots, N$, and $N = 1, 2, \ldots$.

Here $d = 1$, $\mathbf{X} = (1, a^{T/N}, \ldots, a^{(M-1)T/N})^{\mathrm{T}}$ for $M < N$,

$$\mathbf{X}\mathbf{X}^{\mathrm{T}} = \sum_{k=0}^{M-1} a^{2kT/N} = \frac{a^{2MT/N}-1}{a^{2T/N}-1}, \quad \mathbf{X}^{\mathrm{T}} e_M = a^{(M-1)T/N},$$

$$\text{and} \quad \vartheta_M^2 = \vartheta_M^2(N) = (a^{2T/N}-1)\frac{a^{2(M-1)T/N}}{a^{2MT/N}-1} = (1-a^{-2T/N})(1-a^{-2MT/N}).$$

Since $N \to \infty$, it follows that $1 - a^{-2T/N} \sim 2\ln(a)T/N$. We consider the case of $M \to \infty$. If $M/N \to \beta > 0$, then $N\vartheta_M^2(N) \to 2\ln(a)T(1-a^{-2T\beta})$. If $M = o(N)$, then $N\vartheta_M^2(N) \to 0$.

3. *Polynomial series.* Here

$$f_n = \beta_p n^p + \ldots + \beta_1 n + \beta_0, \quad \beta_p \neq 0.$$

In this case, $\dim \mathrm{U}_0^{\perp}(M) = p + 1$ for $M \geq p + 1$, and vectors $P_\ell = Z_M(\ell)/\|Z_M(\ell)\|$ with $Z_M(\ell) = \left(0^\ell, 1^\ell, \ldots, (M-1)^\ell\right)^{\mathrm{T}}$ form a basis in the linear space $\mathrm{U}_0^{\perp}(M)$.

We set $\mathbf{X} = [P_0, \ldots, P_p]$. Then $\|Z_M(\ell)\|^2 \sim M^{2\ell+1}/(2\ell+1)$ and $\mathbf{X}_M^{\mathrm{T}}\mathbf{X}_M \to \mathbf{C}_1$, where $\mathbf{C}_1$ is a $(p+1) \times (p+1)$ positive definite matrix with the entries

$$c_{ij}^{(1)} = \frac{\sqrt{(2i-1)(2j-1)}}{i+j-1}, \quad i, j = 1, \ldots, p+1.$$

Let $A_p := \left(1, \ldots, \sqrt{2\ell+1}, \ldots, \sqrt{2p+1}\right)^{\mathrm{T}}$. Then $\sqrt{M} e_M^{\mathrm{T}}\mathbf{X}_M \to A_p^{\mathrm{T}}$ and $M\vartheta_M^2 \to A_p^{\mathrm{T}}\mathbf{C}_1^{-1}A_p > 0$ as $M \to \infty$.

If $p = 1$, then $M\vartheta_M^2 \to 1/4$ as $M \to \infty$.

4. *Trigonometric series*. We consider the series

$$f_n = \sum_{i=1}^{p} \beta_i \cos(2\pi\omega_i n + \varphi_i) \quad \text{with} \quad \beta_i \neq 0 \quad \text{and} \quad \omega_i \in (0,1/2),$$

where $\omega_j \neq \omega_i$ for $j \neq i$. If $M \geq 2p$, then $\dim \mathbb{U}_0^{\perp}(M) = 2p$.

We set $\cos_{k\ell} = \cos(2\pi k\omega_\ell)$, $\sin_{k\ell} = \sin(2\pi k\omega_\ell)$,

$$C_{jl} = (\cos_{0\ell},\ldots,\cos_{j-1\ell})^{\mathrm{T}}, \quad S_{jl} = (\sin_{0\ell},\ldots,\sin_{j-1\ell})^{\mathrm{T}},$$

and $\mathbf{X} = [C_{M1} : \ldots : C_{Mp} : S_{M1} : \ldots : S_{Mp}]$. Then $M(\mathbf{X}\mathbf{X}^{\mathrm{T}})^{-1}/2 \to \mathbf{I}_{2p}$, where $\mathbf{I}_k$ is the $k \times k$ unit matrix.

Similarly, $\left\|\mathfrak{e}_M^{\mathrm{T}}\mathbf{X}_M\right\|^2 \to 2p$. Therefore, $\vartheta_M^2 \sim 4p/M$ as $M \to \infty$.

We note that if $0 \in \{\omega_1, \ldots, \omega_p\}$ and/or $1/2 \in \{\omega_1, \ldots, \omega_p\}$, then $\vartheta_M^2 \sim c/M$ with some $c > 0$ as before. In particular, $\vartheta_M^2 = 1/M$ for $f_n = c = \mathrm{const} \neq 0$.

2.2. Construction of a LRF

We consider the series $\mathrm{F} = (f_1, \ldots, f_n, \ldots)$ and assume that the linear space $\mathbb{U}_0^{\perp}(M)$ with $\dim \mathbb{U}_0^{\perp}(M) = d$ is not vertical for some $M > 1$. We define $\mathbb{U}_0(M)$ as an orthogonal complement of $\mathbb{U}_0^{\perp}(M)$, and $\mathbf{P}_0 = \mathbf{P}_0(M)$, as an orthogonal projector onto $\mathbb{U}_0(M)$. Moreover, we introduce an $(M-1) \times M$ matrix $\mathbf{G}_M$ by the equality

$$\mathbf{G}_M = \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 0 & 0 \\ 0 & 0 & \ldots & 0 & 1 & 0 \end{pmatrix}.$$

The following proposition in somewhat different notation can be found in ([1], Theorem 5.2 and Proposition 5.3). We present a shorter proof of this fact.

Proposition 1. *We set*

$$R := (b_{M-1},\ldots,b_1)^{\mathrm{T}} = -\frac{1}{1-\vartheta_M^2}\mathbf{G}_M\mathbf{P}_0\mathfrak{e}_M. \tag{6}$$

Then

$$f_n = \sum_{k=1}^{M-1} b_k f_{n-k}, \quad n \geq M. \tag{7}$$

Moreover,

$$\|R\|^2 = \frac{\vartheta_M^2}{1-\vartheta_M^2}. \tag{8}$$

Proof. We note that the vector $\mathbf{P}_0 Y$ is orthogonal to $\mathbb{U}_0^{\perp}(M)$ for any $Y \in \mathbb{R}^M$. Therefore, $(\mathbf{P}_0 Y, F_{n-M+1,n}) = 0$ for $n \geq M$.

We denote by z_M the last coordinate of the vector $\mathbf{P}_0 Y$. If $z_M \neq 0$, then the equality $(\mathbf{P}_0 Y, F_{n-M+1,n}) = 0$ can be rewritten as $(Q, F_{n-M+1,n}) = 0$, where

$$Q = \frac{1}{z_M}\mathbf{P}_0 Y = \begin{pmatrix} -\tilde{R} \\ 1 \end{pmatrix} \quad \text{with} \quad \tilde{R} = -\mathbf{G}_M Q = -\frac{1}{z_M}\mathbf{G}_M\mathbf{P}_0 Y.$$

Since the equalities $(Q, F_{n-M+1,n}) = 0$ and $f_n = (\tilde{R}, F_{n-M+1,n-1})$ are equivalent, we arrive at a LRF of the form (7), which governs the series F. Thus, it remains to prove that the choice of $Y = \mathfrak{e}_M$ leads to $z_M = 1 - \vartheta_M^2$. This follows from formula (1): $(\mathbf{P}_0\mathfrak{e}_M, \mathfrak{e}_M) = \|\mathbf{P}_0\mathfrak{e}_M\|^2 = 1 - \vartheta_M^2$, and equality (7) follows. Using the formula

$$\begin{pmatrix} -R \\ 1 \end{pmatrix} = \frac{1}{1-\vartheta_M^2}\mathbf{P}_0\mathfrak{e}_M,$$

we arrive at (8).

3. RECURRENT FORECAST AND ITS ACCURACY

In ([1], Subsection 2.1), the authors used formula (6) to obtain an approximation of the vector R in the case when the series F is corrupted by additive noise, i.e., the series E.

Let the "signal" $\mathrm{F} = (f_1, \ldots, f_n, \ldots)$ be governed by the minimal LRF (3) of order d. In addition, we consider the "noise" series $\mathrm{E} = (e_1, \ldots, e_n, \ldots)$, set $\mathrm{F}_N = (f_1, \ldots, f_N)$, $\mathrm{E}_N = (e_1, \ldots, e_N)$, and finally $\mathrm{X}_N = (f_1 + \delta e_1, \ldots, f_N + \delta e_N)$, where δ is a formal perturbation parameter. It is assumed that the series X_N is known, and the goal is to forecast the values $f_{N+1}, \ldots, f_{N+S}$ of the series F for some $S \geq 1$.

To explain how this is done, we start by approximating the coefficient vector of the LRF (6).

3.1. Approximating LRF

In ([1], Ch. 2.), the authors described the following method for obtaining an approximating LRF.

1. *Embedding.* After choosing the window length $M < N$, the series X_N is transformed into an $M \times K$ Hankel matrix $\mathbf{H}(\delta)$ with the entries $\mathbf{H}(\delta)[ij] = x_{i+j-1}$, $1 \leq i \leq M$, $1 \leq j \leq K := N - M + 1$. In addition, it is assumed that $\min(M, K) \geq d$.

2. *Singular value decomposition and special grouping.* We seek the best (with respect to the Frobenius norm) approximation $\tilde{\mathbf{H}}(\delta)$ of the matrix $\mathbf{H}(\delta)$ among all $M \times K$ matrices of rank d. This is done by singular value decomposition of the matrix $\mathbf{H}(\delta)$ and summation of the d principal elementary matrices of this decomposition.

We denote the linear space spanned by the columns of the matrix $\tilde{\mathbf{H}}(\delta)$ by $\mathrm{U}_0^{\perp}(\delta)$, and its orthogonal complement, by $\mathrm{U}_0(\delta)$. In addition, we consider $\mathbf{P}_0^{\perp}(\delta)$ and $\mathbf{P}_0(\delta)$, which are orthogonal projections onto these subspaces.

3. If the space $\mathrm{U}_0^{\perp}(\delta)$ is not vertical, then the vector

$$R(\delta) = -\frac{1}{\|\mathbf{P}_0(\delta)\mathfrak{e}_M\|^2}\mathbf{G}_M\mathbf{P}_0(\delta)\mathfrak{e}_M$$

is proposed as an approximation of the vector R introduced in (6).

The following proposition (see ([4], Proposition 5.1)) is used to estimate the accuracy of this approximation.

Proposition 2. *We denote* $\Delta\mathbf{P}(\delta) = \|\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp}\|$, *where* $\|\mathbf{A}\|$ *is the spectral norm of the matrix* $\mathbf{A}$. *If* $\Delta\mathbf{P}(\delta) < \|\mathbf{P}_0\mathfrak{e}_M\|$, *then* $\mathrm{U}_0^{\perp}(\delta)$ *is not vertical and*

$$\|R(\delta) - R\| \leq \frac{\Delta\mathbf{P}(\delta)}{1-\vartheta_M^2}\left(1 - \frac{\Delta\mathbf{P}(\delta)}{\sqrt{1-\vartheta_M^2}}\right)^{-2}\left(1 + \frac{2}{\sqrt{1-\vartheta_M^2}}\right).$$

Remark 1. Proposition 2 gives sufficient conditions for the convergence $\|R(\delta) - R\| \to 0$ in terms of $\|\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp}\|$ and ϑ_M as $M \to \infty$. Namely, if

(1) $\liminf_M \vartheta_M < 1$;

(2) $\|\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp}\| \to 0$ as $M \to \infty$, then $\|R(\delta) - R\| = O\left(\|\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp}\|\right) \to 0$. As follows from the examples in Subsection 2.1, the first condition looks quite natural. As for the second condition, many examples of its implementation are given in ([4], Subsection 3.2).

3.2. Recurrent Forecast

Let us describe how a recurrent forecast is constructed on the basis of the approximation of the LRF considered in the previous section.

Since (see formula (7)), $f_{N+1} = (R, F_{N-M+1,N})$ with $F_{N-M+1,N} = (f_{N-M+1}, \ldots, f_N)^{\mathrm{T}}$, it follows that, to obtain a forecast $\tilde{f}_{N+1}$ of the value f_{N+1} with the help of the approximating LRF $R(\delta)$, one should have a good approximation $\tilde{F}_{N-M+1,N} = (f_{N-M+1}(\delta), \ldots, f_N(\delta))^{\mathrm{T}}$ of the vector $F_{N-M+1,N}$.

Then the number

$$\tilde{f}_{N+1} = (R(\delta), \tilde{F}_{N-M+1,N}) \tag{9}$$

is called the *one-step forecast of the series* F_N. For $k > 1$, the k-step forecast of the series F_N is defined recursively (see [1], Subsection 2.1). For example, $\tilde{f}_{N+2} := (R(\delta), \tilde{F}^*_{N-M+2,N+1})$ with $\tilde{F}^*_{N-M+2,N+1} = (f_{N-M+2}(\delta), \ldots, f_N(\delta), \tilde{f}_{N+1})^{\mathrm{T}}$.

There are many ways to estimate the values f_i of the signal F_N. In the basic version of the SSA forecast, this is done as follows. First of all, one chooses a new *window length* L, which plays exactly the same role as the window length M in Subsection 3.1. We note that only the case of $L = M$ is described in [1] and [2].

Next, the first two points of constructing the approximating LFR (embedding, singular value decomposition, and special grouping) described in Section 3.1 are implemented with M replaced by L. As before, the resulting $L \times K$ matrix with $K = N - L + 1$ is denoted by $\tilde{\mathbf{H}}(\delta)$.

Then a Hankel matrix $\hat{\mathbf{H}}(\delta)$ closest to $\tilde{\mathbf{H}}(\delta)$ in the Frobenius norm is constructed. Finally, applying the inverse of the embedding, we obtain the series $\mathrm{F}_N(\delta) = (f_1(\delta), \ldots, f_N(\delta))$, which is considered as an approximation to F_N.

It is this choice of approximation $\tilde{F}_{N-M+1,N}$ of the vector $F_{N-M+1,N}$ in formula (9) that leads to a one-step recurrent forecast of the series F_N by the SSA method.

3.3. Forecast Accuracy. Basic Formula

Let us now turn to estimating the accuracy of the recurrent forecast. Denoting $\Delta_{N-M+1,N} = \tilde{F}_{N-M+1,N} - F_{N-M+1,N}$ and $\Delta(R) = R(\delta) - R$, we obtain

$$\begin{aligned} \left|\tilde{f}_{N+1} - f_{N+1}\right| &= \left|(R(\delta), \tilde{F}_{N-M+1,N}) - (R, F_{N-M+1,N})\right| \\ &\le \left|(\Delta(R), F_{N-M+1,N})\right| + \left|(R, \Delta_{N-M+1,N})\right| + \left|(\Delta(R), \Delta_{N-M+1,N})\right| \\ &\le \|\Delta(R)\|\|F_{N-M+1,N}\| + \|R\|\|\Delta_{N-M+1,N}\| + \|\Delta(R)\|\|\Delta_{N-M+1,N}\|. \end{aligned} \tag{10}$$

Thus, the upper bound of the one-step forecast accuracy depends on:

(1) the absolute values of the signal f_i for $i = N - M + 1, \ldots, N$;

(2) the norm $\|R\|$ [see (6)]. According to (8), this norm is expressed in terms of the verticality coefficient ϑ_M. In particular, if $\vartheta_M \to 0$ as $M \to \infty$, then $\|R\| \sim \vartheta_M$;

(3) the accuracy $\Delta(R) = R(\delta) - R$ of the approximating LRF. In view of Proposition 2 and Remark 1, it is natural to expect that $\|R(\delta) - R\| = O\left(\left\|\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp}\right\|\right)$. As already mentioned, many examples concerning the convergence rate of $\left\|\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp}\right\|$ to zero can be found in ([4], Subsection 3.2);

(4) the absolute values of the approximation errors $r_i(\delta) = f_i(\delta) - f_i$ for the last M values of the signal F_N.

In this case, the general approach to theoretical estimation of the quantities $|r_i(\delta)|$ as $N \to \infty$ by the SSA was published in [4]. Although this procedure may be quite laborious, below we give some examples of its application.

3.4. Accuracy of Recurrent Forecasting. Examples

As already mentioned, inequality (10) makes it possible to obtain an upper bound for the one-step forecast accuracy for the series F_N. In this section we provide a number of examples where all components of this estimate can be theoretically analyzed for large N, L, and M.

Example 1. *Exponential signal and harmonic noise*.

We consider $f_n = a^n$ with $a > 1$ and $e_n = \cos(2\pi\omega n + \varphi)$ with $\omega \in (0, 1/2)$.

Theoretical results.

1. $\vartheta_M^2 \to 1 - 1/a^2$ as $M \to \infty$ (see Subsection 2.1.1 of this paper).

2. If $M/N \to \beta \in (0, 1)$, then, for any δ, $\Delta\mathbf{P}(\delta) \sim c\sqrt{N}a^{-N}$ with some $c > 0$, see ([4], Subsection 3.2.1)).

3. If $L \le K$ and $L/N \to \alpha \in (0, 1/2]$, then $r_j(\delta) = \rho_j(\delta) + O(N^2 a^{-N})$, where

$$|\rho_j(\delta)| \le C \begin{cases} a^{-(L-j)}/j & \text{for } 1 \le j < L, \\ 1/L & \text{for } L \le j < K, \\ 1/(N-j+1) + a^{-(N-j+1)} & \text{for } K \le j \le N \end{cases} \tag{11}$$

for any δ. This result was published in [6].

Upper bound for $|\tilde{f}_{N+1} - f_{N+1}|$.

We consider the second term $\|R\|\|\Delta_{N-M+1,N}\| = \|R\|\|\tilde{F}_{N-M+1,N} - F_{N-M+1,N}\|$ on the right-hand side of (10). According to (8), $\|R\|^2 = \vartheta_M^2/(1 - \vartheta_M^2) \to \text{const} > 0$ as $M \to \infty$.

From (11) it is easy to find that $\|\Delta_{N-M+1,N}\|$ does not tend to zero as $N \to \infty$. Therefore, under these conditions, the right-hand side of (10) does not tend to zero either. Of course, this does not imply that the left-hand side of (10) does not tend to zero, but computer experiments (see the Appendix) confirm this hypothesis.

Example 2. *Scheme of series for exponential signal and harmonic noise.*

Here $f_n = f_n^{(N)} = a^{nT/N}$ with $a > 1$, $0 < n \le N$, $N \ge 1$, and $e_n = \cos(2\pi\omega n + \varphi)$ with $0 < \omega < 1/2$.

Theoretical results.

1. If $M/N \to \beta \in (0, 1)$, then $N\vartheta_M^2 \to \text{const} > 0$ (see Subsection 2.1.1).

2. Under the same conditions, there exists $\delta_0 > 0$ such that $\Delta\mathbf{P}(\delta) = O(N^{-1})$ for $|\delta| < \delta_0$.

3. If $L/N \to \alpha \in (0, 1)$, then there exists $\delta_0' > 0$ such that $\max_{1\le i\le N} |r_i(\delta)| = O(N^{-1})$ for $|\delta| < \delta_0'$.

The proof of both assertions can be found in [6].

Upper bound for $|\tilde{f}_{N+1} - f_{N+1}|$.

Let $L/N \to \alpha$ and $M/N \to \beta$, $\alpha, \beta \in (0, 1)$.

Since

$$\|\Delta(R)\| = \|R(\delta) - R\| = O\left(\|\mathbf{P}_0^\perp(\delta) - \mathbf{P}_0^\perp\|\right) = O(1/M)$$

and

$$\|F_{N-M+1,N}\|^2 = \sum_{i=N-M+1}^{N} f_i^2 = O(M),$$

the first term $J_1 = \|\Delta(R)\|\|F_{N-M+1,N}\|$ on the right-hand side of (10) has the form $J_1 = O(1/\sqrt{M}) = O(1/\sqrt{N})$. Similarly, $\|R\|^2 \sim \|\mathbf{P}_0^\perp e_M\|^2 = O(1/M)$ and

$$\|\Delta_{N-M+1,N}\|^2 = \|\tilde{F}_{N-M+1,N} - F_{N-M+1,N}\|^2 = \sum_{i=N-M=1}^{N} (f_i(\delta) - f_i)^2 = O(MN^{-2}).$$

Therefore, we obtain the estimate $J_2 = O(N^{-1})$ for the second term $J_2 = \|R\|\|\Delta_{N-M+1,N}\|$ on the right-hand side of (10). Since $\Delta(R) = o(\|R\|)$, it follows that $|\tilde{f}_{N+1} - f_{N+1}| = O(1\sqrt{N})$.

Example 3. *Linear signal and harmonic noise.*

Here $f_n = an + b$ and $e_n = \cos(2\pi\omega n + \varphi)$, $0 < \omega < 1/2$.

Theoretical results.

1. If $M \to \infty$, then $M\vartheta_M^2 \to \text{const} > 0$ (see Subsection 2.1.1).

2. If $N \to \infty$ is odd and $M = (N+1)/2$, then $\Delta\mathbf{P}(\delta) = O(N^{-2})$ for any δ.

3. For odd N and $L = (N+1)/2$, $\max_{1\le i\le N}|r_i(\delta)| = O(N^{-1})$ for any δ.

Both assertions follow from the results published in [7]. General considerations and computer experiments suggest that similar asymptotics for $\Delta\mathbf{P}(\delta)$ and $\max_{1\le i\le N}|r_i(\delta)|$ should remain valid as $M/N \to \beta$ and $L/N \to \alpha$ if $\alpha, \beta \in (0, 1)$.

Upper bound for $|\tilde{f}_{N+1} - f_{N+1}|$.

Since $f_n \sim an$, we have

$$\|F_{N-M+1,N}\|^2 = \sum_{i=N-M+1}^{N} f_i^2 = O(N^3)$$

and $\|F_{N-M+1,N}\| = O(N^{3/2})$ as $N \to \infty$. Since $|f_i(\delta) - f_i| = O(N^{-1})$,

$$\|F_{N-M+1,N}\|^2 = \sum_{i=N-M+1}^{N} (f_i(\delta) - f_i)^2 = O(MN^{-2}) = O(N^{-1})$$

and $\|F_{N-M+1,N}\| = O(N^{-1/2})$.

Further, $\|R(\delta)) - R\| = O(\Delta\mathbf{P}(\delta)) = O(1/N^2)$, and $\|R\|^2 \sim \left\|\mathbf{P}_0^{\perp}e_M\right\|^2 = O(1/M)$, i.e., $\|R\| = O(1/\sqrt{N})$. Therefore,

$$J_1 = \|R(\delta) - R\|\|F_{N-M+1,N}\| = O(N^{-1/2}), \quad J_2 = \|R\|\|\Delta_{N-M+1,N}\| = O(N^{-1}),$$

and, since $\|R(\delta)) - R\| = o(\|R\|)$, the one-step forecasting error is $O(N^{-1/2})$.

Example 4. *Constant signal and sawtooth noise*.

Here $f_n = 1$ and $e_n = (-1)^n$.

Theoretical results.

1. If $M \to \infty$, then $M\vartheta_M^2 \to \text{const} > 0$ (see Subsection 2.1.1).

2. If $M/N \to \beta \in (0, 1)$, then there exists a $\delta_0 > 0$ such that $\Delta\mathbf{P}(\delta) = O(N^{-1})$ for $|\delta| < \delta_0$.

3. If $L/N \to \alpha \in (0, 1)$, then there exists a $\delta_0' > 0$ such that $\max_{1\le i\le N}|r_i(\delta)| = O(N^{-1})$ for $|\delta| < \delta_0'$. Both results can be found in ([4], Subsection 5.3.1).

Computer experiments give hope that similar results will hold for $f_n = \cos(2\pi\omega_1 n + \varphi_1)$ and $e_n = \cos(2\pi\omega_2 n + \varphi_2)$, where $\omega_1, \omega_2 \in (0, 1/2)$ and $\omega_1 \neq \omega_2$.

Upper bound for $|\tilde{f}_{N+1} - f_{N+1}|$.

Since all theoretical results for this case are analogous to those already described in Example 2, we have $|\overline{f}_{N+1} - f_{N=1}| = O(1/\sqrt{N})$.

APPENDIX. COMPUTER EXPERIMENTS

In this section we present several variants of computer experiments that illustrate the theoretical results of Subsection 3.4. As already mentioned, the considered series have the form $x_n = f_n + \delta e_n$, $1 \le n \le N$, and the problem lies in forecasting the signal value f_{N+k}, $k \ge 1$.

We consider three variants of the signal f_n:

1. Exponential signal (EXP): $f_n = a^n$, $a = 1.01$, and $\delta = 1$.

2. Linear signal (LIN): $f_n = an + b$, $a = 0.5$, $b = 1$, and $\delta = 0.5$.

3. Harmonic signal (COS): $f_n = \cos(2\pi\omega_0 n)$ with $\omega_0 = \sqrt{2}/2$, and $\delta = 0.5$.

In all cases, the noise has the form $e_n = \cos(2\pi\omega n)$ with $\omega = \sqrt{3}/2$.

As before, $\tilde{f}_{N+k}$ denotes the value of the recurrent forecast of f_{N+k}, and $\Delta_k^{(f)}(N) = |\tilde{f}_{N+k} - f_{N+k}|$ is the k-step forecast error.

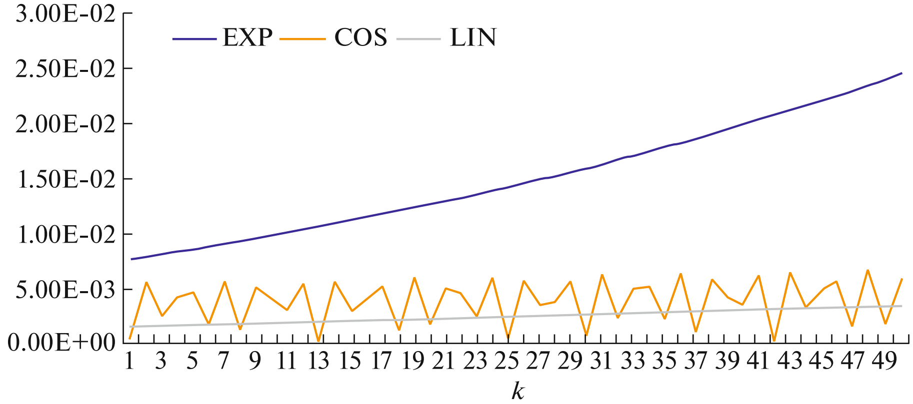

Fig. 1. k-Step forecast errors for exponential (EXP), harmonic (COS) and linear (LIN) signals as a function of k, $k = 1(1)50$.

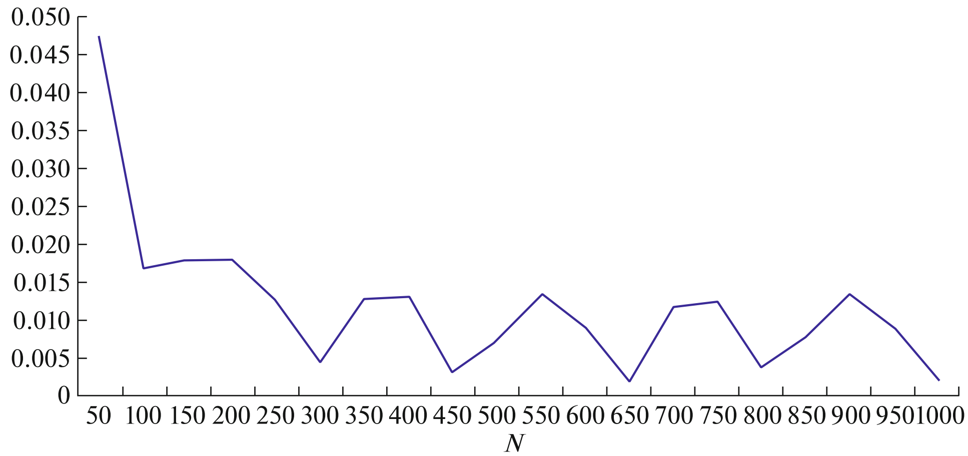

Fig. 2. One-step forecast errors as a function of the series length N for the EXP signal.

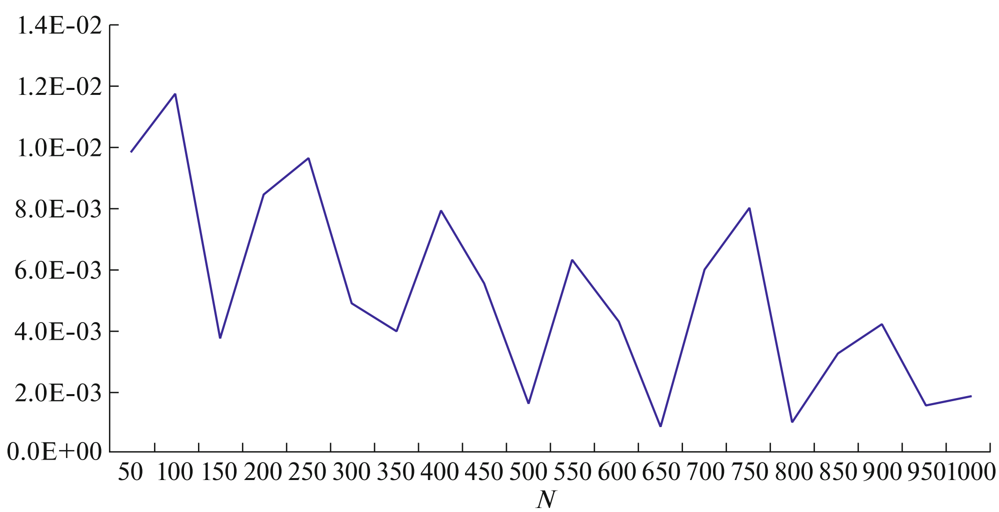

Fig. 3. One-step forecast errors as a function of the series length N for the LIN signal.

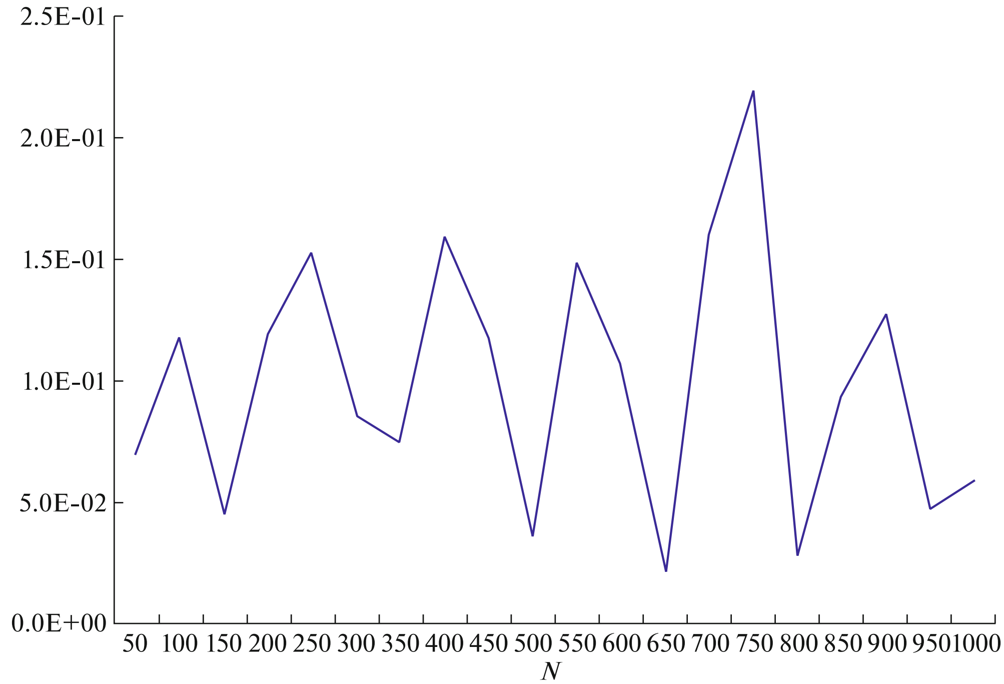

Fig. 4. $\sqrt{N}$ multiplied by one-step forecast errors as a function of the series length N for the LIN signal.

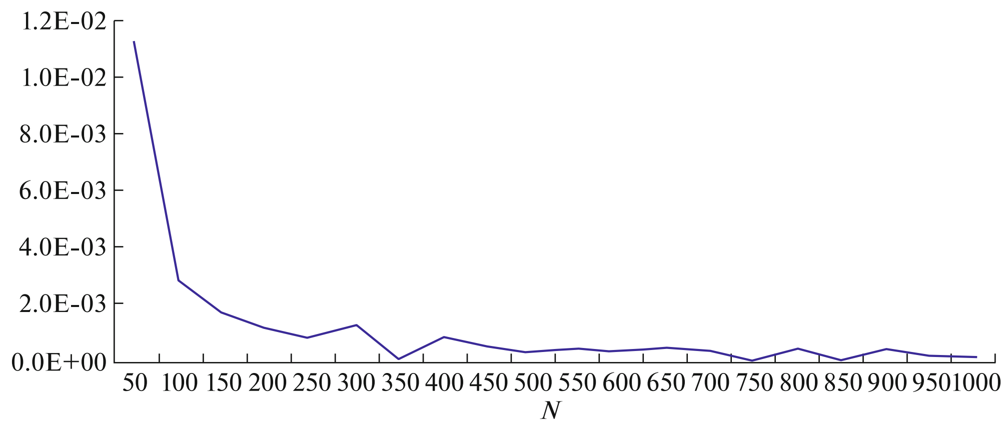

Fig. 5. Mean of the errors for $k = 1(1)10$ forecast steps as a function of the series length N for the COS signal.

k-**Step forecast,** $1 \le k \le 50$**.** Figure 1 is of preliminary character. It shows k-step forecasting errors for exponential (EXP), harmonic (COS), and linear (LIN) signals for $N = 500$, $L = M = 250$, and $k = 1(1)50$. We note that these errors exhibit very smooth behavior for EXP and LIN, while $\Delta_k^{(f)}(N)$ strongly oscillates for COS.

Therefore, below we illustrate the behavior of $\Delta_1^{(f)}(N)$ as a function of N for exponential and linear signals, and take the characteristic $\left(\sum_{k=1}^{10} \Delta_k^{(f)}(N)\right)/10$ to illustrate the forecasting behavior for a harmonic signal.

We note that in all the following experiments, we use the series lengths $N = 50(50)1000$ and $L = M = N/2$.

Exponential signal. Figure 2 shows the behavior of one-step forecasting errors as a function of the series length N for an exponential signal and harmonic noise.

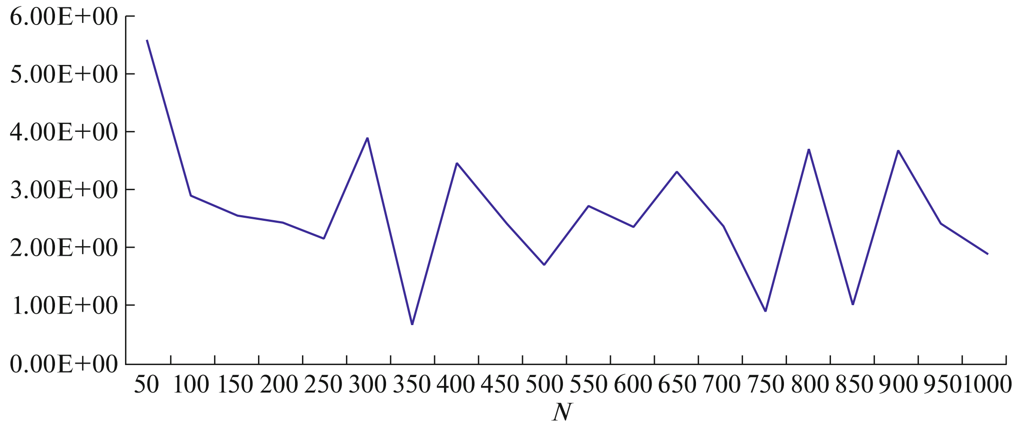

Fig. 6. N multiplied by the mean of the errors for $k = 1(1)10$ forecast steps as a function of the series length N for the COS signal.

The theoretical results of Section 3.4 show that $\Delta_1^{(f)}(N) = O(1)$. Judging by Fig. 2, this estimate is quite accurate: starting from about $N = 300$, the values of $\Delta_1^{(f)}(N)$ do not show an obvious tendency toward a decrease.

Linear signal. As already mentioned, for a linear signal, the expected estimate for $\Delta_1^{(f)}(N)$ is $O(N^{-1/2})$. Figures 3 and 4 confirm this theoretical result.

Harmonic signal. According to the results of Subsection 3.4, the expected upper bound for one-step forecast errors for a harmonic signal is $O(N^{-1/2})$. Figures 5 and 6 show that the actual convergence rate may be $O(N^{-1})$.

ACKNOWLEDGMENTS

I am grateful to the reviewers, whose remarks undoubtedly contributed to improvement of the manuscript.

FUNDING

This work was supported by the Russian Foundation for Basic Research, project no. 20-01-00067.

REFERENCES

1. N. Golyandina, V. Nekrutkin, and A. Zhigljavsky, *Analysis of Time Series Structure. SSA and Related Techniques* (Chapman & Hall/CRC, New York, 2001).
2. N. Golyandina and A. Zhigljavsky, *Singular Spectrum Analysis for Time Series*, 2nd ed. (Springer-Verlag, Berlin, 2020), in Ser.: *Springer Briefs in Statistics*.
3. P. P. Vaidyanathan, *The Theory of Linear Prediction* (Morgan & Claypool, San Rafael, Calif., 2008).
4. V. Nekrutkin, "Perturbation expansions of signal subspaces for long signals," Stat. Its Interface **3**, 297–319 (2010).
5. N. Golyandina, "On the choice of parameters in Singular Spectrum Analysis and related subspace-based methods," Stat. Its Interface **3**, 259–279 (2010).
6. E. Ivanova and V. Nekrutkin, "Two asymptotic approaches for the exponential signal and harmonic noise in Singular Spectrum Analysis," Stat. Its Interface **12**, 49–59 (2019). https://doi.org/10.4310/SII.2019.v12.n1.a5
7. N. V. Zenkova and V. V. Nekrutkin, "On the asymptotical separation of linear signals from harmonics by Singular Spectrum Analysis," Vestn. St Petersburg Univ.: Math. **55**, 166–173 (2022). https://doi.org/10.1134/S1063454122020157

Translated by I. Nikitin