Statistics and Its Interface Volume 10 (2017) 27–32

Asymptotic extraction of common signal subspaces from perturbed signals

Vladimir Nekrutkin* and Irina Vasilinetc

Signal subspaces extracted by Singular Value Decomposition of signal matrices are used in many methods of signal processing. General approach to asymptotic proximity of unperturbed and perturbed signal subspaces is discussed in Nekrutkin 2010, SII, v. 3, 297–319. These theoretical results are illustrated by several examples related to one-dimensional signals and the corresponding Hankel matrices.

In this paper we apply this approach to the multidimensional signals and block-Hankel matrices. More precisely, we suppose that each coordinate of a multidimensional signal produces the same signal subspace. For such signals, we suggest the solution for asymptotic extraction of this subspace from the perturbed multidimensional signal series.

A similar procedure is already used in atmosphere sciences to incorporate both spatial and temporal correlations in multidimensional data.

AMS 2000 subject classifications: Primary 65G99, 65F30; secondary 65F15.

Keywords and phrases: Signal subspace methods, Perturbation expansion, Asymptotic analysis, Multidimensional signals.

1. INTRODUCTION

The general scheme of the signal-subspace approach, studied in [1], can be explained as follows. Consider a one-dimensional or multidimensional "signal" $\mathrm{F}_N = (x_0, \ldots, x_{N-1})$. This series is linearly transformed into $L\times K$ "signal matrix" $\mathbf{H}$. We suppose that the signal F_N and the transformation $\mathrm{F}_N \mapsto \mathbf{H}$ are such that $d \stackrel{\mathrm{def}}{=} \operatorname{rank}\mathbf{H} < \min(L,K)$. Then the linear space $\mathbb{U}_0^{\perp}$ spanned by the columns of matrix $\mathbf{H}$ contains important information about the series F_N. $\mathbb{U}_0^{\perp}$ is further referred to as *signal subspace*.

Assume that we observe the perturbed series $\mathrm{F}_N(\delta) = \mathrm{F}_N+\delta\mathrm{E}_N$, where $\mathrm{E}_N = (e_0, \ldots, e_{N-1})$ is a "noise" series and δ stands for a formal perturbation parameter. Thus, instead of the "signal matrix" $\mathbf{H}$ we work with the perturbed matrix $\mathbf{H}(\delta) = \mathbf{H} + \delta\mathbf{E}$, where the "noise matrix" $\mathbf{E}$ is constructed from the series E_N in the same manner as $\mathbf{H}$ is built from the series F_N.

Consider the Singular Value Decomposition (briefly, SVD) of $\mathbf{H}(\delta)$. If δ is small, then continuity considerations show that the linear space $\mathbb{U}_0^{\perp}(\delta)$ spanned by the left singular vectors for the d largest singular values of this SVD can serve as an approximation of $\mathbb{U}_0^{\perp}$.

Though general results on the proximity of $\mathbb{U}_0^{\perp}$ and $\mathbb{U}_0^{\perp}(\delta)$ are already proved in [1, sect. 2], all examples in [1] are dedicated to one-dimensional signals F_N and Hankel matrices $\mathbf{H}$, $\mathbf{H}(\delta)$.

These examples were of the same structure. We consider an infinite one-dimensional real-valued signal series $\mathrm{F} = (x_0, x_1, \ldots, x_n, \ldots)$ which is governed by a minimal linear recurrent formula of order d. Then finite segments $\mathrm{F}_N = (x_0, x_1, \ldots, x_{N-1})$ of the series F are transformed into signal $L\times K$ Hankel matrices

$$\mathbf{H} = \mathbf{H}_N = \begin{pmatrix} x_0 & x_1 & \ldots & x_{K-1} \\ x_1 & x_2 & \ldots & x_K \\ \vdots & \vdots & \ddots & \vdots \\ x_{L-1} & x_L & \ldots & x_{N-1} \end{pmatrix}, \tag{1.1}$$

where $N\to\infty$, L depends on N and $K = N - L + 1$. All signal matrices have the same rank d in assumption that L and K are sufficiently big. Linear d-dimensional spaces, spanned by columns of matrices $\mathbf{H}$ stand for $\mathbb{U}_0^{\perp} = \mathbb{U}_0^{\perp}(N)$.

The similar procedure is accomplished for the perturbed series $\mathrm{F}(\delta) = \mathrm{F} + \delta\mathrm{E}$ where E stands for some noise series. The result of this transformation is perturbed Hankel matrix $\mathbf{H}(\delta)$ and the linear space $\mathbb{U}_0^{\perp}(\delta)$. The proximity of $\mathbb{U}_0^{\perp}$ and $\mathbb{U}_0^{\perp}(\delta)$ is studied with the help of the corresponding projection operators $\mathbf{P}_0^{\perp}$ and $\mathbf{P}_0^{\perp}(\delta)$ using classical perturbation results [2].

In this paper we consider multidimensional signals with equal signal subspaces for all coordinate series. Each coordinate signal is perturbed by some additive error series. Then the resulting multidimensional series is transformed into the corresponding block-Hankel matrix. Singular Value Decomposition gives us approximate basis of the common signal subspace.

As in [1], the precision of approximation is measured in terms of the sine of the largest principal angle between the perturbed and unperturbed signal subspaces.

Note that similar computational procedure is widely used in atmospheric science to incorporate both the spatial and the temporal correlation in data (see the original paper [4] and the review [5] for details).

*Corresponding author.

Just as in the paper [1], we consider our results as a step to the theoretical foundation of some subspace-based methods, such as MSSA in the style of [6] or different methods for the the parameter estimation of multidimensional signals (see, for example, [7], where some of these methods are briefly described).

2. MAXIMAL AND MINIMAL POSITIVE EIGENVALUES

In the same way as for one-dimensional series, the asymptotic behavior of positive eigenvalues of large matrices play the essential role in the whole consideration. In this section we present several general results on this problem as well as examples related to block-Hankel matrices. The examples are used in Section 3 devoted to main results of the paper.

2.1 General statements

Let us start with two simple assertions formulated as lemmas for convenience of references.

Consider a matrix $\mathbf{G} : \mathbb{R}^K \mapsto \mathbb{R}^L$ and denote $d = \operatorname{rank} \mathbf{G}$. Suppose that

$$\mathbf{G} = \sum_{k=1}^{d} P_k Q_k^{\mathrm{T}} \tag{2.1}$$

with some $P_k \in \mathbb{R}^L$ and $Q_k \in \mathbb{R}^K$, where the vectors $P_1, \ldots, P_d$ (and the vectors $Q_1, \ldots, Q_d$) are linearly independent. Let $\|Z\|$ stand for the Euclidean norm of the vector Z. Denote $X_i = P_i/\|P_i\|$, $Y_i = Q_i/\|Q_i\|$,

$$\mathbf{X} = [X_1 : \ldots : X_d], \quad \mathbf{Y} = [Y_1 : \ldots : Y_d], \tag{2.2}$$

$\mathbf{U} = [P_1 : \ldots : P_d]$, and $\mathbf{V} = [Q_1 : \ldots : Q_d]$. Also, set

$$\Pi_P = \begin{pmatrix} \|P_1\| & 0 & \ldots & 0 \\ 0 & \|P_2\| & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \|P_d\| \end{pmatrix},$$

$$\Pi_Q = \begin{pmatrix} \|Q_1\| & 0 & \ldots & 0 \\ 0 & \|Q_2\| & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \|Q_d\| \end{pmatrix},$$

and $\Pi_{PQ} = \Pi_P \Pi_Q$. Lastly, denote

$$\mathbf{C} = \mathbf{X}^{\mathrm{T}}\mathbf{X}\Pi_{PQ}\mathbf{Y}^{\mathrm{T}}\mathbf{Y}\Pi_{PQ} \tag{2.3}$$

and $\mathbf{C}' = \mathbf{U}^{\mathrm{T}}\mathbf{U}\mathbf{V}^{\mathrm{T}}\mathbf{V}$.

Lemma 2.1. *Let λ be a positive eigenvalue of the matrix $\mathbf{G}\mathbf{G}^{\mathrm{T}}$, and let λ correspond to an eigenvector Z. Then*
1. λ is the eigenvalue of the matrix $\mathbf{C}$ and λ corresponds to the (non-null) eigenvector $\mathbf{X}^{\mathrm{T}}Z$.
2. λ is the eigenvalue of the matrix $\mathbf{C}'$ and λ corresponds to the (non-null) eigenvector $\mathbf{U}^{\mathrm{T}}Z$.

Proof. 1. Note that $\mathbf{G} = \mathbf{X}\Pi_{PQ}\mathbf{Y}^{\mathrm{T}}$ and $\mathbf{G}\mathbf{G}^{\mathrm{T}}Z = \mathbf{X}\Pi_{PQ}\,\mathbf{Y}^{\mathrm{T}}\mathbf{Y}\Pi_{PQ}\,\mathbf{X}^{\mathrm{T}}Z = \lambda Z$. Therefore,

$$\lambda\mathbf{X}^{\mathrm{T}}Z = \mathbf{X}^{\mathrm{T}}\mathbf{G}\mathbf{G}^{\mathrm{T}}Z = \big(\mathbf{X}^{\mathrm{T}}\mathbf{X}\Pi_{pq}\mathbf{Y}^{\mathrm{T}}\mathbf{Y}\Pi_{pq}\big)\mathbf{X}^{\mathrm{T}}Z = \mathbf{C}\mathbf{X}^{\mathrm{T}}Z.$$

Since $\mathbf{G}\mathbf{G}^{\mathrm{T}}Z \neq \mathbf{0}$, then $\mathbf{X}^{\mathrm{T}}Z \neq \mathbf{0}$.
2. In view of the equality $\mathbf{G} = \mathbf{U}\mathbf{V}^{\mathrm{T}}$, $\mathbf{G}\mathbf{G}^{\mathrm{T}}Z = \mathbf{U}\mathbf{V}^{\mathrm{T}}\mathbf{V}\mathbf{U}^{\mathrm{T}}Z$ and

$$\lambda\mathbf{U}^{\mathrm{T}}Z = \mathbf{U}^{\mathrm{T}}\mathbf{G}\mathbf{G}^{\mathrm{T}}Z = \big(\mathbf{U}^{\mathrm{T}}\mathbf{U}\mathbf{V}^{\mathrm{T}}\mathbf{V}\big)\mathbf{U}^{\mathrm{T}}Z = \mathbf{C}'\mathbf{U}^{\mathrm{T}}Z.$$

Since $\mathbf{G}\mathbf{G}^{\mathrm{T}}Z \neq \mathbf{0}$, then $\mathbf{U}^{\mathrm{T}}Z \neq \mathbf{0}$. □

Corollary 2.1. *The set of positive eigenvalues of the $L\times L$ matrix $\mathbf{G}\mathbf{G}^{\mathrm{T}}$ coincides with the spectrum of the $d \times d$ matrices $\mathbf{C}$ and $\mathbf{C}'$.*

Now consider a sequence of $d \times d$ matrices $\mathbf{C}_n$ such that all eigenvalues of matrices $\mathbf{C}_n$ are positive for any n and denote by $\lambda_{\max}^{(n)}$ and $\lambda_{\min}^{(n)}$ maximal and minimal eigenvalues of the matrix $\mathbf{C}_n$.

Lemma 2.2. *1. Assume that $a_n \to +\infty$ and $\mathbf{C}_n/a_n \to \mathbf{M}_1$. If the matrix $\mathbf{M}_1$ is not nilpotent, then $\lambda_{\max}^{(n)}/a_n \to \theta_1 > 0$, where θ_1 is the maximal eigenvalue of the matrix $\mathbf{M}_1$.*
2. Assume that $b_n \to +\infty$ and $b_n\mathbf{C}_n^{-1} \to \mathbf{M}_2$. If the matrix $\mathbf{M}_2$ is not nilpotent, then $\lambda_{\min}^{(n)}/b_n \to 1/\theta_2 > 0$, where θ_2 is the maximal eigenvalue of the matrix $\mathbf{M}_2$.

Proof. 1. The continuity considerations show that all eigenvalues of the matrix $\mathbf{M}_1$ are real and non-negative. Since $\mathbf{M}_1$ is not nilpotent, then some of them are positive. Now the assertion follows from the fact that $X^{\mathrm{T}}\mathbf{C}_n X/a_n \to X^{\mathrm{T}}\mathbf{M}_1 X$ uniformly on the set $\{X \in \mathbb{R}^d$ such that $\|X\| = 1\}$.
2. The second assertion is proved in the same manner taking in consideration that $\lambda_{\min}^{(n)}$ is the maximal eigenvalue of the matrix $\mathbf{C}_n^{-1}$. □

The next proposition is the particular case of Lemma 2.2 for matrices discussed in Lemma 2.1.

Consider a sequence of $L_n \times K_n$ matrices $\mathbf{G} = \mathbf{G}_n$ with fixed $d = \operatorname{rank} \mathbf{G}$. As in Lemma 2.1 we use representations $\mathbf{G} = \sum_{k=1}^{d} P_k Q_k^{\mathrm{T}}$ with some $P_k \in \mathbb{R}^L$ and $Q_k \in \mathbb{R}^K$, where the vectors $P_1, \ldots, P_d$ (and the vectors $Q_1, \ldots, Q_d$) are linearly independent. (Vectors P_k and Q_k depend on n, still we ignore this dependence in our notation.) Let σ_j stand for $\|P_j\|\|Q_j\|$. It is convenient to assume that $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_{d-1} \geq \sigma_d$.

Define matrices $\mathbf{X} = \mathbf{X}^{(n)}$, $\mathbf{Y} = \mathbf{Y}^{(n)}$, $\Pi_P = \Pi_P^{(n)}$, $\Pi_Q = \Pi_Q^{(n)}$, $\Pi_{PQ} = \Pi_{PQ}^{(n)}$, and $\mathbf{C} = \mathbf{C}_n$ as earlier. Note that for any n, matrices $\mathbf{C}_n$ have the fixed size $d \times d$.

Lastly, denote by $\lambda_{\max}$ and $\lambda_{\min}$ maximal and minimal positive eigenvalues of the matrix $\mathbf{G}\mathbf{G}^{\mathrm{T}}$.

Proposition 2.1. *Assume that $\mathbf{X}^{\mathrm{T}}\mathbf{X} \to \mathbf{A}$ and $\mathbf{Y}^{\mathrm{T}}\mathbf{Y} \to \mathbf{B}$, where both $\mathbf{A}$ and $\mathbf{B}$ are invertible.*
1. Assume that $\sigma_1 \to \infty$ and $\sigma_j/\sigma_1 \to c_j$ for any j. Set

$$\Lambda_{\max} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & c_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & c_d \end{pmatrix}$$

and

$$\mathbf{M}_1 = \mathbf{A}\Lambda_{\max}\mathbf{B}\Lambda_{\max}. \tag{2.4}$$

If the matrix $\mathbf{M}_1$ is not nilpotent, then $\lambda_{\max}/\sigma_1^2 \to \theta_1 > 0$, where θ_1 is the maximal eigenvalue of the matrix $\mathbf{M}_1$.
2. Assume that $\sigma_d \to \infty$ and $\sigma_d/\sigma_j \to c'_j$ for any j. Set

$$\Lambda_{\min} = \begin{pmatrix} c'_1 & 0 & \dots & 0 \\ 0 & c'_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}$$

and

$$\mathbf{M}_2 = \Lambda_{\min}\mathbf{B}^{-1}\Lambda_{\min}\mathbf{A}^{-1}. \tag{2.5}$$

If the matrix $\mathbf{M}_2$ is not nilpotent, then $\lambda_{\min}/\sigma_d^2 \to 1/\theta_2 > 0$, where θ_2 is the maximal eigenvalue of the matrix $\mathbf{M}_2$.

Proof. 1. By Corollary 2.1, the set of positive eigenvalues of the matrix $\mathbf{G}\mathbf{G}^{\mathrm{T}}$ coincides with the spectrum of the $d \times d$ matrix $\mathbf{C} = \mathbf{X}^{\mathrm{T}}\mathbf{X}\Pi_{PQ}\mathbf{Y}^{\mathrm{T}}\mathbf{Y}\Pi_{PQ}$.

Since $\mathbf{X}^{\mathrm{T}}\mathbf{X} \to \mathbf{A}$, $\mathbf{Y}^{\mathrm{T}}\mathbf{Y} \to \mathbf{B}$ and

$$\frac{1}{\sigma_1}\Pi_{PQ} = \frac{1}{\sigma_1}\begin{pmatrix} \sigma_1 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma_d \end{pmatrix} \to \Lambda_{\max},$$

then $\mathbf{C}/\sigma_1^2 \to \mathbf{M}_1$. Therefore the result follows from the first assertion of Lemma 2.2.
2. Note that matrices $\mathbf{X}^{\mathrm{T}}\mathbf{X}$ and $\mathbf{Y}^{\mathrm{T}}\mathbf{Y}$ are strictly positive definite and hence invertible. This means that

$$\mathbf{C}^{-1} = \Pi_{PQ}^{-1}\big(\mathbf{Y}^{\mathrm{T}}\mathbf{Y}\big)^{-1}\Pi_{PQ}^{-1}\big(\mathbf{X}^{\mathrm{T}}\mathbf{X}\big)^{-1}.$$

Since matrices $\mathbf{A}$ and $\mathbf{B}$ are of full rank and

$$\sigma_d\,\Pi_{PQ}^{-1} = \sigma_d\begin{pmatrix} 1/\sigma_1 & 0 & \dots & 0 \\ 0 & 1/\sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1/\sigma_d \end{pmatrix} \to \Lambda_{\min},$$

then $\sigma_d^2\,\mathbf{C}^{-1} \to \mathbf{M}_2$. The second assertion of Lemma 2.2 finishes the proof. □

Corollary 2.2. *The following remark helps to verify the conditions of Proposition 2.1. Assume that $\mathbf{X}^{\mathrm{T}}\mathbf{X} \to \mathbf{A}$ and $\mathbf{Y}^{\mathrm{T}}\mathbf{Y} \to \mathbf{B}$, where both $\mathbf{A}$ and $\mathbf{B}$ are invertible. Let $\mathbf{A} = \{a_{ij}\}_{i,j=1}^d$, $\mathbf{A}^{-1} = \{\alpha_{ij}\}_{i,j=1}^d$, $\mathbf{B} = \{b_{ij}\}_{i,j=1}^d$ and $\mathbf{B}^{-1} = \{\beta_{ij}\}_{i,j=1}^d$. Since $\mathbf{A}$ and $\mathbf{B}$ are strictly positive definite, the diagonal elements of these matrices are positive.*
1. If all diagonal elements of the matrix $\Lambda_{\max}$ are positive, then $\det \mathbf{M}_1 \neq 0$ and $\lambda_{\min}/\sigma_1^2$ tends to the minimal eigenvalue of the matrix $\mathbf{M}_1$.
2. If all but the first diagonal elements of the matrix $\Lambda_{\max}$ are equal to zero, then the matrix $\mathbf{C}_1$ is not nilpotent and $\lambda_{\max}/\sigma_1 \to a_{11}b_{11}$.
3. If all but the last diagonal elements of the matrix $\Lambda_{\min}$ are equal to zero, then the matrix $\mathbf{C}_2$ is not nilpotent and $\lambda_{\min}/\sigma_d \to 1/(\alpha_{dd}\,\beta_{dd})$.

Proof. The first statement of the corollary immediately follows from the first assertion of Proposition 2.1. The two remaining statements are similar. Let us demonstrate the first of them. Describe matrices $\mathbf{A}$ and $\mathbf{B}$ by their columns:

$$\mathbf{A} = \big[A_1 : \dots : A_d\big], \quad \mathbf{B} = \big[B_1 : \dots : B_d\big].$$

Then

$$\mathbf{C}_1 = \mathbf{A}\Lambda_{\max}\mathbf{B}\Lambda_{\max} = \big[A_1 : 0 : \dots : 0\big]\big[B_1 : 0 : \dots : 0\big] = \big[C_1 : 0 : \dots : 0\big]$$

with $C_1 = \big(a_{11}b_{11}, \dots, a_{d1}\,b_{11}\big)^{\mathrm{T}}$. The latter matrix is not nilpotent and its maximal eigenvalue equals $a_{11}b_{11}$. □

2.2 Examples: eigenvalues of some block-Hankel matrices

Consider the two-dimensional series of length N with components $x_n^{(1)}$ and $x_n^{(2)}$, $0 \le n \le N-1$. For a certain $L \in \{2, \dots, N\}$, let $\mathbf{G}^{(i)}$ be the "trajectory" (Hankel) matrix

$$\mathbf{G}^{(i)} = \begin{pmatrix} x_0^{(i)} & x_1^{(i)} & \dots & x_{K-1}^{(i)} \\ x_1^{(i)} & x_2^{(i)} & \dots & x_K^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ x_{L-1}^{(i)} & x_L^{(i)} & \dots & x_{K+L-2}^{(i)} \end{pmatrix} \tag{2.6}$$

of the series $x_n^{(i)}$ and put $\mathbf{G} = [\mathbf{G}^{(1)} : \mathbf{G}^{(2)}]$. Lastly, let $\lambda_{\max}$ and $\lambda_{\min}$ stand for the maximum and minimum positive eigenvalues of the matrix $\mathbf{G}\mathbf{G}^{\mathrm{T}}$.

Lemma 2.3. *1. Consider the two-dimensional series with components*

$$x_n^{(1)} = \sum_{m=1}^{r_1}\alpha_m\cos(2\pi\omega_m n), \; x_n^{(2)} = \sum_{k=1}^{r_2}\beta_k\cos(2\pi\nu_k n), \tag{2.7}$$

where $\omega_m, \nu_k \in (0, 1/2)$, $\alpha_m, \beta_k \neq 0$, $\omega_i \neq \omega_j$ and $\nu_i \neq \nu_j$ for $i \neq j$.

Then there exist $\Lambda_{\max} \ge \Lambda_{\min} > 0$ such that

$$\lambda_{\max}/LK \to \Lambda_{\max}, \quad \lambda_{\min}/LK \to \Lambda_{\min}$$

as $N \to \infty$ and $\min(L, K) \to \infty$.

2. Consider two polynomial series

$$x_n^{(1)} = \sum_{j=0}^{p_1} \gamma_j n^j, \quad x_n^{(2)} = \sum_{m=0}^{p_2} \beta_m n^m \tag{2.8}$$

with $\gamma_{p_1}, \beta_{p_2} \neq 0$ *and* $p_1 \geq p_2$. *If* $L/N \to \alpha \in (0,1)$ *then there exist* $\theta_{\max} \geq \theta_{\min} > 0$ *such that*

$$N^{-2p_1-2}\lambda_{\max} \to \theta_{\max}, \quad N^{-2p_1-2}\lambda_{\min} \to \theta_{\min}.$$

3. Consider exponential signals with components

$$x_n^{(1)} = \sum_{\ell=1}^{p} \alpha_\ell a_\ell^n, \quad x_n^{(2)} = \sum_{\ell=1}^{p} \beta_\ell a_\ell^n, \tag{2.9}$$

where $a_1 > \ldots > a_p > 1$, $\alpha_i, \beta_i \neq 0$. *Then* $\lambda_{\max}/a_1^{2N}$ *and* $\lambda_{\min}/a_p^{2N}$ *tend to positive constants as* $\min(L,K) \to \infty$.

Proof. The demonstrations of all assertions are similar. To avoid elementary but laborious calculations we present only the general line of the proof.

First of all, we express the matrix $\mathbf{G}$ in the form (2.1) with the appropriate vectors P_k and Q_k. Note that the rank d of this matrix does not depend on the length N of the series under consideration provided that N is sufficiently big.

Note that $d = 2\,\mathrm{card}\,(\mathcal{N}_1 \cup \mathcal{N}_2)$ with $\mathcal{N}_1 = \{\omega_1, \ldots, \omega_{r_1}\}$ and $\mathcal{N}_2 = \{\nu_1, \ldots, \nu_{r_2}\}$ for the series (2.7), $d = p_1 + 1$ for the series (2.8) and $d = p$ for the series (2.9).

Then we apply the first assertion of Lemma 2.1 with Corollary 2.1 and therefore reduce our eigenvalue problem from the $L \times L$ matrix $\mathbf{G}\mathbf{G}^{\mathrm{T}}$ to the $d \times d$ matrix (2.3). To find the asymptotic behavior of maximal and minimal eigenvalues of the matrix (2.3) as $N \to \infty$, we use Lemma 2.2 and Proposition 2.1 with its Corollary 2.2.

More precisely, it can be checked that matrices $\mathbf{X}\mathbf{X}^{\mathrm{T}}$ and $\mathbf{Y}\mathbf{Y}^{\mathrm{T}}$ (see (2.2) for their definition) tend to invertible matrices $\mathbf{A}$ and $\mathbf{B}$ as $N \to \infty$. Moreover, matrices (2.4) and (2.5) are not nilpotent. Thus we can use Proposition 2.1. By this proposition, $\lambda_{\max} \sim \max_j \|P_j\|\,\|Q_j\|$ and $\lambda_{\min} \sim \min_j \|P_j\|\,\|Q_j\|$ as $N \to \infty$. □

Lemma 2.4. *Let*

$$e_n^{(1)} = \sum_{j=-\infty}^{\infty} b_j\, \epsilon_{j+n}^{(1)}, \quad e_n^{(2)} = \sum_{j=-\infty}^{\infty} c_j\, \epsilon_{j+n}^{(2)}, \tag{2.10}$$

where $\epsilon_n^{(i)}$ *are i.i.d. random variables, defined on the same probability space* $(\Omega, \mathfrak{F}, \mathbb{P})$. *Assume that* $\mathbb{E}\epsilon_n^{(i)} = 0$, $\mathbb{D}\epsilon_n^{(i)} = 1$, *and* $\mathbb{E}|\epsilon_n^{(i)}|^3 < \infty$.

Additionally, it is supposed that $\sum_j |b_j| < \infty$, $\sum_j b_j^2 = 1$, $\sum_j |c_j| < \infty$, *and* $\sum_j c_j^2 = 1$.

If $\max(L,K) \to \infty$, *then there exist* $\Omega' \in \mathfrak{F}$ *with* $\mathbb{P}(\Omega_0) = 1$ *and a constant* C *such that for any* $\omega \in \Omega_0$

$$\limsup_N \frac{\|\mathbf{G}\|}{N \ln N} \leq C.$$

Proof. It can be easily seen that the maximal singular value of the matrix $(\mathbf{G}^{(1)} : \mathbf{0})$ is equal to the maximal singular value of $\mathbf{G}^{(1)}$. Analogously, $\|(\mathbf{0} : \mathbf{G}^{(2)})\| = \|\mathbf{G}^{(2)}\|$.

Applying [3, th. 1] we see that almost surely

$$\limsup_N \frac{\|\mathbf{G}^{(i)}\|}{N \ln N} \leq C_i$$

for $i = 1,2$ and certain constants C_1, C_2.

Since $\mathbf{G} = (\mathbf{G}^{(1)} : \mathbf{0}) + (\mathbf{0} : \mathbf{G}^{(2)})$ then

$$\limsup_N \frac{\|\mathbf{G}\|}{N \ln N} \leq \limsup_N \frac{\|\mathbf{G}^{(1)}\| + \|\mathbf{G}^{(2)}\|}{N \ln N} \leq C = C_1 + C_2$$

with probability 1. □

Remark 2.1. It is worth mentioning that the random sequences $\{\epsilon_n^{(1)}\}$ and $\{\epsilon_n^{(2)}\}$ in (2.10) are not assumed to be independent.

Remark 2.2. Though both lemmas 2.3 and 2.4 are formulated for two-dimensional series, the analogous results hold for m-dimensional series with $m > 2$.

3. ON THE COMMON PART OF SEVERAL SIGNAL SUBSPACES

For $1 \leq i \leq m$ let us consider m real-valued series $F_N^{(i)} = (x_0^{(i)}, x_1^{(i)}, \ldots, x_{N-1}^{(i)})$ and the corresponding Hankel matrices $\mathbf{H}_i$ defined by the right-hand side of (2.6).

All series together are described by the block-Hankel matrix $\mathbf{H} = (\mathbf{H}_1 : \ldots : \mathbf{H}_m)$.

If each of series $F_N^{(i)}$ is governed by a certain (minimal) LRF of order d_i and if $L = L(N)$ and K are big enough, then the linear space $\mathbb{U}_0^{\perp}$ spanned by the columns of matrix $\mathbf{H}$ has some fixed dimension d for any L and K.

Note that here Hankel matrices $\mathbf{H}_1, \ldots, \mathbf{H}_m$ are put side by side to form the matrix $\mathbf{H}$ and we use their columns to define $\mathbb{U}_0^{\perp}$. This corresponds to the general construction described in [1]. The other possible way to build $\mathbf{H}$ is to stack $\mathbf{H}_i$ one on the top of the other. Then $\mathbb{U}_0^{\perp}$ must be spanned by the rows of the matrix $\mathbf{H}$.

We call $\mathbb{U}_0^{\perp}$ *the signal subspace of the multidimensional signal* $(F_N^{(1)}, \ldots, F_N^{(m)})$. Of course, $\mathbb{U}_0^{\perp}$ is the smallest linear space containing all signal subspaces $\mathbb{U}_i^{\perp}$ of one-dimensional series $F_N^{(i)}$.

In the same manner as it is described in the Introduction, each signal series $F_N^{(i)}$ is perturbed with the help of some "error" series $E_N^{(i)}$, and the resulting perturbed series $F_N^{(i)}(\delta) = F_N^{(i)} + \delta E_N^{(i)}$ are transformed into the block-Hankel matrix

$$\begin{aligned}\mathbf{H}(\delta) &= \big(\mathbf{H}_1(\delta) : \ldots : \mathbf{H}_m(\delta)\big)\\ &= \big(\mathbf{H}_1 + \delta\mathbf{E}_1 : \ldots : \mathbf{H}_m + \delta\mathbf{E}_m\big) = \mathbf{H} + \delta\mathbf{E}.\end{aligned}$$

Taking d leading left singular vectors of the matrix $\mathbf{H}(\delta)$, we obtain the perturbed signal subspace $\mathbb{U}_0^{\perp}(\delta)$ and our general aim is to investigate the behavior of the main principal angle $\|\mathbf{P}_0^\perp(\delta)-\mathbf{P}_0^\perp\|$ between $\mathbb{U}_0^\perp(\delta)$ and $\mathbb{U}_0^\perp$ as $N\to\infty$.

Here we restrict ourselves to the case when all signal subspaces $\mathbb{U}_i^\perp$ coincide. This case can be interpreted as the problem of extracting the common part of signal subspaces of the observed series $F_N^{(i)}(\delta)$. For the sake of brevity, we take $m=2$.

More precisely, we consider 3 kinds of two-dimensional signals:

1) oscillating signals with components

$$x_n^{(1)}=\sum_{i=1}^{r}\alpha_i\cos\big(2\pi\omega_i n\big),\ \ x_n^{(2)}=\sum_{i=1}^{r}\beta_i\cos\big(2\pi\omega_i n\big), \tag{3.1}$$

where $\omega_l\in(0,1/2)$, $\omega_i\neq\omega_j$ for $i\neq j$, $|\alpha_i|>0$ and $|\beta_i|>0$;

2) polynomial signals with components

$$x_n^{(1)}=\sum_{l=0}^{p}\alpha_l n^l,\quad x_n^{(2)}=\sum_{l=0}^{p}\beta_l n^l, \tag{3.2}$$

where $p\geq 1$ and $\beta_p,\alpha_p\neq 0$, and

3) exponential signals with components defined by (2.9), where $a_1>\ldots>a_p>1$, $\alpha_i\neq 0$, $\beta_i\neq 0$.

Proposition 3.1. *Consider the two-dimensional signal with components defined by* (3.1) *and assume that* $\min(L,K)\to\infty$ *as* $N\to\infty$.

1. Suppose that the "error series" has components

$$e_n^{(1)}=\sum_{i=1}^{r_1}b_i\cos\big(2\pi\nu_i^{(1)}n\big),\quad e_n^{(2)}=\sum_{i=1}^{r_2}c_i\cos\big(2\pi\nu_i^{(2)}n\big),$$

with $\nu_i^{(1)},\nu_i^{(2)}\in(0,1/2)$ *and* $b_i\neq 0$, $c_i\neq 0$.

If the sets $\{\nu_i^{(1)},\nu_j^{(2)}\}$ *and* $\{\omega_i\}$ *are disjoint, then*

$$\|\mathbf{P}_0^\perp(\delta)-\mathbf{P}_0^\perp\|=|\delta|O\big(1/\min(L,K)\big)$$

for any δ *such that* $|\delta|<\delta_0=\delta_0(b_i,c_i,\alpha_i,\beta_i)$ *and for any* $N\geq N_0(\delta)$.

2. If the error components are defined by (2.10), *then* $\|\mathbf{P}_0^\perp(\delta)-\mathbf{P}_0^\perp\|=|\delta|O\big(\sqrt{N\ln N}/\sqrt{LK}\big)$ *for any* δ *with probability 1.*

Proof. 1. Let $\mu_{\min},\mu_{\max}$ be minimal and maximal positive eigenvalues of the matrix $\mathbf{HH}^{\mathrm{T}}$ and set $\nu_{\max}=\|\mathbf{EE}^{\mathrm{T}}\|$. Denote

$$\Theta_1=\sqrt{\frac{\nu_{\max}}{\mu_{\max}}},\quad \Theta_2=\frac{\mu_{\max}}{\mu_{\min}},\quad \Theta=\Theta_1\Theta_2,$$

and $\Delta=1/\limsup_N(\Theta\Theta_1)$. Then $\Delta>0$ in view of the first assertion of Lemma 2.3. Set

$$C_{Li}=\big(1,\cos(\pi\omega_i),\ldots,\cos((L-1)\pi\omega_i)\big)^T,$$
$$S_{Li}=\big(0,\sin(\pi\omega_i),\ldots,\sin((L-1)\pi\omega_i)\big)^T.$$

If μ is a positive eigenvalue of $\mathbf{HH}^{\mathrm{T}}$, then each eigenvector $U_{\mu,j}$, corresponding to this eigenvalue has the form

$$U_{\mu,j}=\sum_{i,j}\Big(a_{Li}^{(\mu,j)}C_{Li}+b_{Li}^{(\mu,j)}S_{Li}\Big),$$

where $1\leq j\leq d_\mu$, d_μ is the multiplicity of μ, and the number of pairs (i,j) equals $2r$.

We specify coefficients $a_{Li}^{(\mu,j)},b_{Li}^{(\mu,j)}$ by conditions that $\|U_{\mu,j}\|=1$ and $U_{\mu_1,j_1}^{\mathrm{T}}U_{\mu_2,j_2}=0$ for different pairs (μ_1,j_1), (μ_2,j_2). Then $\mathbf{P}_\mu=\sum_j U_{\mu,j}U_{\mu,j}^{\mathrm{T}}$, where $\mathbf{P}_\mu$ is the projection operator on the eigenspace corresponding to positive eigenvalue μ of the matrix $\mathbf{HH}^{\mathrm{T}}$.

It is easy to check that uniformly in L

$$\|U_{\mu_1,j_1}U_{\mu_2,j_2}^{\mathrm{T}}\mathbf{E}_i\mathbf{E}_i^{\mathrm{T}}\|=O(K)$$

for $i=1,2$ and any pairs (μ_1,j_1), (μ_2,j_2).

Since $\mathbf{EE}^{\mathrm{T}}=\mathbf{E}_1\mathbf{E}_1^{\mathrm{T}}+\mathbf{E}_2\mathbf{E}_2^{\mathrm{T}}$, then uniformly in L

$$\|U_{\mu_1,j_1}U_{\mu_2,j_2}^{\mathrm{T}}\mathbf{E}_i\mathbf{E}_i^{\mathrm{T}}\|=O(K) \tag{3.3}$$

also for $i=1,2$ and any pairs (μ_1,j_1), (μ_2,j_2).

Let $\mathbf{S}_0=\sum_{\mu>0}\mathbf{P}_\mu/\mu$ be the pseudoinverse to $\mathbf{HH}^{\mathrm{T}}$. Since all μ have the order $O(LK)$ as $L,K\to\infty$, then $\|\mathbf{S}_0\mathbf{EE}^{\mathrm{T}}\|=O(1/L)$ in view of (3.3).

Lastly, direct calculations show that $\|\mathbf{HE}^{\mathrm{T}}\|=O(L)$. Therefore, we can apply [1, Proposition 3.2]. By this proposition,

$$\|\mathbf{P}_0^\perp(\delta)-\mathbf{P}_0^\perp\|=|\delta|O\Big(\big\|\mathbf{HE}^{\mathrm{T}}\big\|/\mu_{\min}+|\delta|\,\big\|\mathbf{S}_0\mathbf{EE}^{\mathrm{T}}\big\|\Big)$$

for $\delta<\delta_0=\Delta/4$. As it was already noticed, $\mu_{\min}$ has the order LK as $L,K\to\infty$, and the first assertion is proved.

2. To prove the second assertion, we use both lemmas 2.3 and 2.4 as well as [1, Proposition 3.1]. This proposition affirms that under condition $\Theta\to 0$ as $N\to\infty$,

$$\limsup_N \Theta^{-1}\|\mathbf{P}_0^\perp(\delta)-\mathbf{P}_0^\perp\|\leq 8C|\delta|$$

for any δ and some constant $C>0$. In view of Lemma 2.3, $\Theta_2=\mu_{\max}/\mu_{\min}\to\Lambda_{\max}/\Lambda_{\min}>0$, while (see Lemma 2.4) $\Theta_1=\sqrt{\nu_{\max}/\mu_{\max}}\leq\sqrt{N\ln N/LK}$ with probability 1. Therefore, the second assertion is also proved. □

Proposition 3.2. *Consider the two-dimensional signal with components defined by* (3.2) *and assume that* $L/N\to\alpha\in(0,1)$ *as* $N\to\infty$.

1. If the error series is determined by (2.7), *then* $\|\mathbf{P}_0^\perp(\delta)-\mathbf{P}_0^\perp\|=|\delta|O\big(N^{-p}\big)$ *for any* δ.

2. If the error components are defined by (2.10), *then* $\|\mathbf{P}_0^\perp(\delta)-\mathbf{P}_0^\perp\|=|\delta|O\big(\sqrt{\ln N}N^{-p-1/2}\big)$ *with probability 1 for any* δ.

Proof. In view of the first assertion of Lemma 2.3, $\nu_{\max}/LK$ tends to some positive constant. On the other hand, the second assertion of the same Lemma 2.3 shows, that $\mu_{\max}\asymp$

N^{2p+2} and $\mu_{\min} \asymp N^{2p+2}$. Therefore,

$$\Theta = \sqrt{\frac{\nu_{\max}}{\mu_{\max}} \frac{\mu_{\max}}{\mu_{\min}}} \asymp \sqrt{LK} N^{-p-1} \asymp N^{-p} \to 0$$

as $N \to \infty$. Applying [1, proposition 3.1], we see that

$$\limsup_N \Theta^{-1} \|\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp}\| = O\big(|\delta|\big), \tag{3.4}$$

and the proof is complete.

2. The second assertion is proved in the same manner taking into account that $\nu_{\max} = O(\sqrt{N \ln N})$ almost surely in view of Lemma 2.4. □

Proposition 3.3. *Consider the two-dimensional exponential signal with components defined by (2.9) under the restriction $\tau \stackrel{\text{def}}{=} a_1/a_p^2 < 1$. Assume that $\min(L, K) \to \infty$.*

1. If the error series is defined by (2.7), then

$$\|\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp}\| = |\delta| O\big(\sqrt{LK}\, \tau^N\big)$$

for any δ as $N \to \infty$.

2. If the error components are defined by (2.10), then

$$\|\mathbf{P}_0^{\perp}(\delta) - \mathbf{P}_0^{\perp}\| = |\delta| O\big(N\sqrt{\ln N}\, \tau^N\big)$$

with probability 1 for any δ.

Proof. 1. In view of the third assertion of Lemma 2.3, $\mu_{\max} \asymp a_1^{2N}$ and $\mu_{\min} \asymp a_p^{2N}$. The first assertion of the same lemma tells that $\nu_{\max} \asymp LK$. Thus $\Theta \asymp \sqrt{LK}\, \tau^N$ and (3.4) gives us the result.

2. The second assertion is proved in the same manner taking into account that $\nu_{\max} = O(\sqrt{N \ln N})$ almost surely in view of Lemma 2.4. □

ACKNOWLEDGEMENTS

The author is grateful to both anonymous referees for many useful comments and suggestions which helped to improve the paper.

Received 24 December 2014

REFERENCES

[1] Nekrutkin, V. (2010). Perturbation expansions of signal subspaces for long signals. *Statistics and Its Interface.* **3** 297–319. MR2720134

[2] Kato, T. (1966). *Perturbation Theory for Linear Operators.* Springer-Verlag, Berlin-Heidelberg-New York. MR0203473

[3] Nekrutkin, V. (2013). Remark on the norm of random Hankel matrices. *Vestnik St. Petersburg University. Mathematics.* **46** 4 189–192. MR3184394

[4] Weare, B. C. and Nasstrom J. S. (1982). Examples of extended empirical orthogonal functions. *Monthly Weather Review.* **110** 481–485.

[5] Hannachi, A., Jolliffe, I. T., and Stephenson, D. B. (2007). Empirical orthogonal functions and related techniques in atmospheric science: A review. *Int. J. Climatol.* **27** 1119–1152.

[6] Golyandina, N. and Stepanov, D. (2005). SSA-Based approaches to analysis and forecast of multidimensional time series. *Proceedings of the 5th St. Petersburg Workshop on Simulation, June 26–July 2, 2005, St. Petersburg State University, St. Petersburg*, 293–298. URL http://www.gistatgroup.com/gus/ mssa2.pdf.

[7] Sahnoun, S., Djermoune E-H., and David Brie, D. (2010). A comparative study of subspace-based methods for 2D nuclear magnetic resonance spectroscopy signals. URL https://hal.archives-ouvertes.fr/hal-00627648v2.

Vladimir Nekrutkin
Math. Department, St. Petersburg State University
Universitetsky av. 28, 198504
St. Petersburg, Petrodvorets
Russia
E-mail address: vnekr@statmod.ru

Irina Vasilinetc
Algorithmic Biology Lab, St. Petersburg Academic University
8/3 Khlopina Str, 194021
St. Petersburg
Russia
E-mail address: vasilinetc.ira@gmail.com