NONPARAMETRIC STATISTICS FOR SUBSPACE BASED FREQUENCY ESTIMATION

S. Visuri[1]* H. Oja[2] V. Koivunen[1]
[1]Signal Processing Laboratory
Helsinki Univ. of Technology
P.O. Box 3000, FIN-02015 HUT, FINLAND
{samuli.visuri,visa.koivunen}@hut.fi
[2]Dept. of Statistics
University of Jyväskylä
P.O. Box 35, FIN-40351 Jyväskylä, FINLAND
hannu.oja@jyu.fi

ABSTRACT

The paper introduces new subspace based frequency estimation methods. The techniques are based on estimating the noise or signal subspace from the sample spatial sign autocovariance matrix. The theoretical motivation for the techniques is shown under the white Gaussian noise assumption. A simulation study is performed to demonstrate the robust performance of the algorithms both in Gaussian and non-Gaussian noise. The results imply that when the noise is Gaussian, the proposed methods have similar good performance as the standard subspace methods (MUSIC, ESPRIT). When the noise is heavy-tailed, the proposed methods outperform the standard subspace techniques.

1 INTRODUCTION

Many subspace methods for frequency estimation have been developed in the signal processing research community (see e.g., [1]). Most of these techniques employ the autocovariance matrix of the assumed signal model and in particular its eigendecomposition, in finding the frequencies. In the algorithms, the autocovariance matrix is estimated using the sample autocovariance matrix. The sample autocovariance matrix is an accurate estimator if the noise is assumed to be white Gaussian distributed.

A drastic degradation in performance is experienced when the noise is non-Gaussian. It is well known that in non-Gaussian noise, the sample autocovariance matrix may give strongly misleading estimation results. The perturbations caused by such noise may change the eigenvalue spectrum and the directions of the eigenvectors significantly.

This paper introduces new robust subspace frequency estimation methods. These methods are highly efficient, i.e., they have almost optimal performance at nominal conditions (Gaussian noise). Moreover, they work reliably in wide range of different noise conditions including the case of heavy-tailed noise where the second order moments do not exist. The proposed techniques stem from multivariate nonparametric statistics.

The paper is organized as follows. In section 2, we define the theoretical concepts from multivariate nonparametric statistics needed in the paper. The signal model and the theoretical basis for subspace algorithms are introduced in section 3. New algorithms for frequency estimation are proposed in section 4 and section 5 deals with the estimation of the number of complex exponentials. The simulation results are given in section 6. Finally, section 7 concludes the paper.

*Financial support for this work was provided by the Academy of Finland

2 SPATIAL SIGN AUTOCOVARIANCE MATRIX

We begin with some definitions. For a complex $M \times 1$ vector $\mathbf{y}$, the spatial sign function $\mathbf{S}$ is defined as

$$\mathbf{S}(\mathbf{y}) = \begin{cases} \frac{\mathbf{y}}{||\mathbf{y}||}, & \mathbf{y} \neq \mathbf{0} \\ \mathbf{0}, & \mathbf{y} = \mathbf{0}, \end{cases}$$

where $||\mathbf{y}|| = (\mathbf{y}^H\mathbf{y})^{1/2}$. For a time series $x_1, \ldots, x_N$, the sample spatial Sign Autocovariance Matrix (SAM) of size $M \times M$ is

$$R_S = \frac{1}{N-M+1} \sum_{i=1}^{N-M+1} \mathbf{S}(\mathbf{z}_i)\mathbf{S}^H(\mathbf{z}_i), \qquad (1)$$

where

$$\mathbf{z}_i = [x_i, \ldots, x_{i+M-1}]^T, \quad i = 1, \ldots, N-M+1.$$

Note that the usual estimate for the $M \times M$ autocovariance matrix is the sample covariance matrix of the $\mathbf{z}_i$s.

3 SUBSPACE FREQUENCY ESTIMATION

In this paper we deal with the following complex exponentials in white noise model:

$$x(n) = \sum_{i=1}^{p} A_i e^{jn\omega_i} + v(n), \qquad (2)$$

where ω_i, $i = 1, \ldots, p$, are the frequencies ($\omega_l \neq \omega_k$ for $k \neq l$) and $v(n)$ is the complex valued circular white noise. The complex amplitudes A_i are given by

$$A_i = |A_i| e^{j\phi_i},$$

where ϕ_i is the phase.

Let $\mathbf{z}(k) = [x(k), \ldots, x(k+M-1)]^T$, $M > p$, $\mathbf{a} = [A_1, \ldots, A_p]^T$ and $\mathbf{v}(k) = [v(k), \ldots, v(k+M-1)]^T$. We can now write

$$\mathbf{z}(k) = BD(k)\mathbf{a} + \mathbf{v}(k),$$

where

$$B = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{j\omega_1} & e^{j\omega_2} & \ldots & e^{j\omega_p} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j(M-1)\omega_1} & e^{j(M-1)\omega_2} & \ldots & e^{j(M-1)\omega_p} \end{bmatrix}$$

and $D(k) = \mathrm{diag}[e^{jk\omega_1}, \ldots, e^{jk\omega_p}]$. If the initial phases ϕ_i are independent and uniformly distributed on $[-\pi, \pi]$, the $M \times M$ autocovariance matrix of $x(n)$ is

$$\Sigma = E\{\mathbf{z}(k)\mathbf{z}(k)^H\} = B\Sigma_{\mathbf{a}}B^H + \sigma^2 I. \qquad (3)$$

where σ^2 is the noise variance and $\Sigma_a = \mathrm{diag}\{|A_1|^2, \ldots, |A_p|^2\}$. It then follows that the $M - p$ smallest eigenvalues of the matrix Σ are equal to the noise variance σ^2 and the corresponding eigenvectors are orthogonal to the columns of the matrix B. These eigenvectors span the *noise subspace* and the eigenvectors corresponding to the p largest eigenvalues span the *signal subspace*.

The subspace frequency estimation algorithms are then based on the different properties of these two subspaces. The MUSIC algorithm, for example, uses the orthogonality property. In this algorithm, the frequency estimates are chosen to be the p largest peaks in the pseudospectrum

$$P_M(\omega) = \frac{1}{\mathbf{b}^H(\omega)\hat{V}\hat{V}^H\mathbf{b}(\omega)},$$

where $\mathbf{b}(\omega) = [1, e^{j\omega}, \ldots, e^{j(M-1)\omega}]^T$ and $\hat{V} = [\hat{\mathbf{v}}_{p+1}, \ldots, \hat{\mathbf{v}}_M]$ is the matrix of the eigenvectors corresponding to the $M-p$ smallest eigenvalues of the sample autocovariance matrix.

We now prove, assuming Gaussian noise, that the signal and noise subspaces can be convergently estimated using the SAM of the observations. The obtained estimates for the signal and noise subspace basis vectors, i.e. the eigenvectors of the SAM, can therefore be used in any subspace frequency estimation method.

Theorem 1 *Assume $\{x(1), \ldots, x(N)\}$ ($N > p$) distributed as given in (2) with deterministic or stochastic initial phases $\phi_1, ..., \phi_p$ and assume the noise v(n) to be i.i.d. Gaussian and independent of the phases. Denote the $M \times M$ ($M > p$) sample SAM of the data by R_S. Then:*

(i) The $M-p$ smallest eigenvalues of $E\{R_S\}$ are equal and the corresponding eigenvectors are orthogonal to the columns of the matrix B.

(ii) As $N \to \infty$,

$$R_S - E\{R_S\} \overset{w.p.1}{\to} 0.$$

Proof. See the Appendix.

The efficiency and robust performance of the SAM based subspace techniques, also in non-Gaussian noise, is shown using simulations in section 6.

4 SUBSPACE ALGORITHMS

We are now ready to give two algorithms illustrating the usage of the SAM in frequency estimation. The first algorithm is a MUSIC-type noise subspace algorithm

SAM-MUSIC

1. Calculate the R_S in (1) of the size $M \times M$ for the data $x_1, \ldots, x_N$.
2. Choose the frequency estimates to be the p highest peaks in the pseudospectrum

$$P_S(\omega) = \frac{1}{\mathbf{b}^H(\omega)\hat{V}_S\hat{V}_S^H\mathbf{b}(\omega)},$$

where $\hat{V}_S$ is the matrix of the eigenvectors of R_S corresponding to the $M - p$ smallest eigenvalues and $\mathbf{b}(\omega)$ is given above.

The second algorithm is based on estimating the signal subspace from the SAM.

SAM-TLS-ESPRIT

1. Calculate the R_S in (1) of the size $M \times M$ for the data $x_1, \ldots, x_N$. Set $\hat{S}$ to be the $M \times p$ matrix of the eigenvectors of R_S corresponding to the p largest eigenvalues.
2. Calculate the total least squares estimate $\hat{\phi}$ for

$$\hat{S}_1\hat{\phi} \approx \hat{S}_2,$$

where $\hat{S}_1 = [I_{M-1}\ \mathbf{0}]\hat{S}$ and $\hat{S}_2 = [\mathbf{0}\ I_{M-1}]\hat{S}$.

3. The frequency estimates are $-\arg(\hat{\nu}_k)$, where $\hat{\nu}_1, \ldots, \hat{\nu}_p$ are the eigenvalues of $\hat{\phi}$.

5 ESTIMATING THE NUMBER OF COMPLEX EXPONENTIALS

If the number of the complex exponentials in (2) is not known, it has to be estimated from the data. A popular method for this kind of tasks is a method based on the Minimum Description Length (MDL) principle. In the MDL based approach, the estimate for the number of

complex exponentials is an integer $k \in \{0, 1, \ldots, M-1\}$ which minimizes the criterion

$$MDL(k) = -\log \left(\frac{\left(\prod_{i=k+1}^{M} \hat{\lambda}_i \right)^{1/(M-k)}}{\frac{1}{M-k} \sum_{i=k+1}^{M} \hat{\lambda}_i} \right)^{(M-k)N} + \frac{1}{2} k(2M-k) \log N, \quad (4)$$

where $\hat{\lambda}_i, i = 1, \ldots, M$, are the eigenvalues of the sample autocovariance matrix. The above criterion was introduced for number of signals estimation in array signal processing [2]. The first term of the criterion compares the equality of the eigenvalues related to the noise subspace eigenvectors and the second term is a penalty term.

It follows from Theorem 1 that when the noise is Gaussian, the noise subspace eigenvalues of R_S will be equal with probability 1. Therefore, we can estimate the number of the complex exponentials also by using the eigenvalues of the R_S and the above criterion. In section 6, we show that this method performs reliably in Gaussian and non-Gaussian noise.

6 SIMULATION RESULTS

In this section, we compare the behavior of the proposed algorithms to conventional methods. The noise model considered is the family of complex isotropic symmetric α-stable ($S\alpha S$) distributions [3]. The characteristic function of $S\alpha S$ distribution is

$$\rho(\omega) = \exp(-\gamma |\omega|^{\alpha}).$$

The smaller the characteristic exponent $\alpha \in [0, 2]$, the heavier the tails of the density (the case $\alpha = 2$ corresponds to Gaussian distribution). The positive valued scalar γ is the dispersion of the distribution. The dispersion plays a role analogous to that of the variance for second order processes.

In our simulations, we use the following signal model:

$$x(n) = \sum_{i=1}^{4} \sqrt{50} e^{\omega_i n + \phi_i} + v(n),$$

where $\omega_1 = 91/72\pi$, $\omega_2 = 89/72\pi$, $\omega_3 = 14/18\pi$, $\omega_4 = 13/18\pi$ and ϕ_i, $i = 1, \ldots, 4$ are uniformly distributed on $[-\pi, \pi]$. The number of observations is $N = 300$ and the size of the matrices (number of lags) used in the algorithms is $M = 30$. In all experiments, the value for the dispersion is $\gamma = 1$.

We first compare the MUSIC and SAM-MUSIC algorithms. Figure 1 shows five realizations of the pseudospectrums for the cases $\alpha = 2$ and $\alpha = 1$. As can be seen from this figure, the estimation methods perform equally in the Gaussian noise ($\alpha = 2$). When the noise is heavy-tailed ($\alpha = 1$), SAM-MUSIC finds the correct frequency peaks reliably, whereas MUSIC often fails.

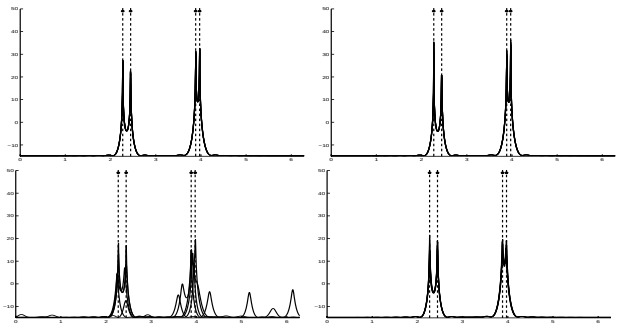

Figure 1: Five realizations of the frequency estimation results for α-stable noise. First column: MUSIC. Second column: SAM-MUSIC. First row: $\alpha = 2$. Second row: $\alpha = 1$. The number of observations is $N = 300$ and the number of lags is $M = 30$.

Figure 2 shows histograms for the estimation results obtained from 200 Monte-Carlo realizations using TLS-ESPRIT and SAM-TLS-ESPRIT algorithms. Similarly to the previous simulation, the performance of the two methods is almost identical when the noise is Gaussian. When the noise is non-Gaussian ($\alpha = 1$), the SAM-TLS-ESPRIT estimates the frequencies significantly better than the TLS-ESPRIT.

In the final simulation we compare the two methods for number of complex exponentials estimation introduced in section 5, when α is varied from 1 to 2. The performance criterion is the relative proportion of correct estimation results and the number of Monte-Carlo realizations is 500. The results are presented in Figure 3. The method based on the sample autocovariance matrix performs poorly for small values of α. On the other hand, the characteristic exponent does not seem to have a significant influence to the estimation method based on the SAM.

7 CONCLUSION

In the paper we propose new subspace methods for frequency estimation. The simulation results imply that the proposed methods perform reliably regardless of the noise distribution, whereas the conventional subspace methods are sensitive to the deviations from Gaussian noise. At the end, it seems worthwhile to mention that the calculation of the SAM is straightforward and therefore the methods require approximately same computational load as the conventional subspace methods.

References

[1] P. Stoica and R. L. Moses, *Introduction to Spectral Analysis*. New York: Wiley, 1997.

[2] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Transactions on Acoustics Speech and Signal Processing*, vol. 33, pp. 387–392, 1985.

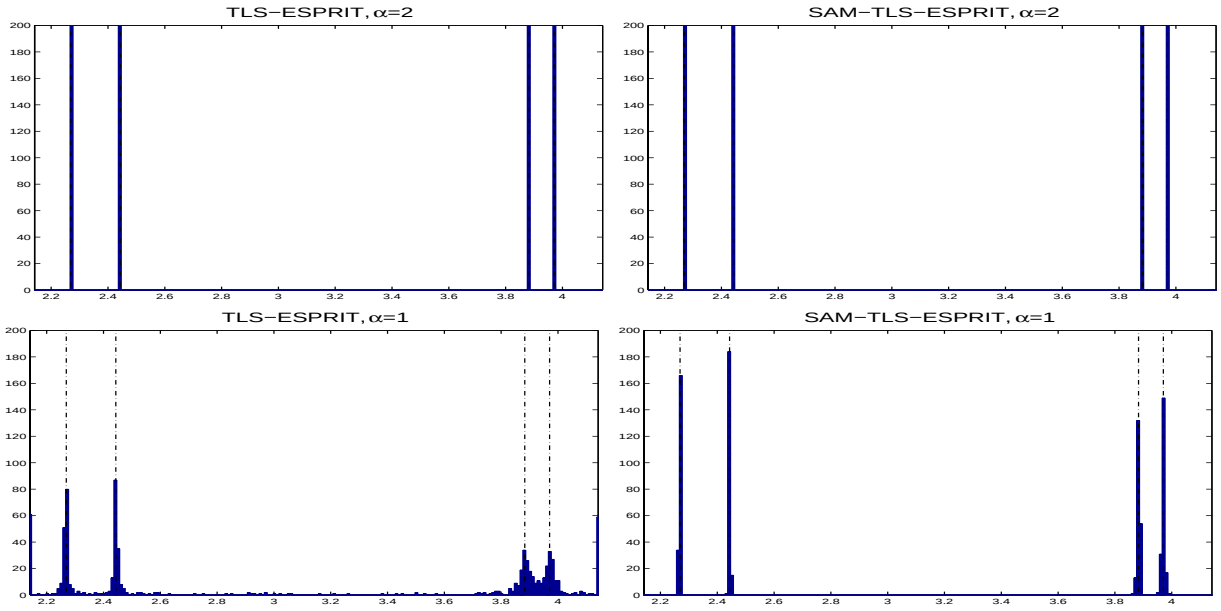

Figure 2: Histograms of the estimation results from 200 Monte-Carlo realizations for TLS-ESPRIT and SAM-TLS-ESPRIT algorithms. The number of observations is $N = 300$ and the number of lags is $M = 30$.

[3] P. Tsakalides and C. L. Nikias, "The robust covariation-based music (ROC-MUSIC) algorithm for bearing estimation in impulsive noise environments," *IEEE Transactions on Signal Processing*, vol. 44, no. 7, pp. 1623–1633, 1996.

[4] S. Visuri, H. Oja, and V. Koivunen, "Direction of arrival estimation based on nonparametric statistics," submitted to the *IEEE Transactions on Signal Processing*, 2000.

[5] R. Serfling, *Approximation Theorems of Mathematical Statistics*. New York: Wiley, 1980.

APPENDIX: PROOF OF THEOREM 1

For proving the result we need the following Lemma:

Lemma 1 *Assume* $\mathbf{x}(n) = G\mathbf{s}(n) + \mathbf{v}(n)$, $n = 1, \ldots, N$, *where* G *is a* $M \times p$ $(M > p)$ *matrix with full column rank. Moreover, assume that the deterministic* p*-vectors* $\mathbf{s}(1), \ldots, \mathbf{s}(N)$ *span a* p*-dimensional subspace and the distribution of the* M*-vector* $\mathbf{v}(n)$ *is complex spherically symmetric* $\forall n$. *Then the* $M - p$ *smallest eigenvalues of* $E\{N^{-1}\sum_{n=1}^{N}\mathbf{S}(\mathbf{x}(n))\mathbf{S}^H(\mathbf{x}(n))\}$ *are equal and the corresponding eigenvectors are orthogonal to the columns of the matrix* G.

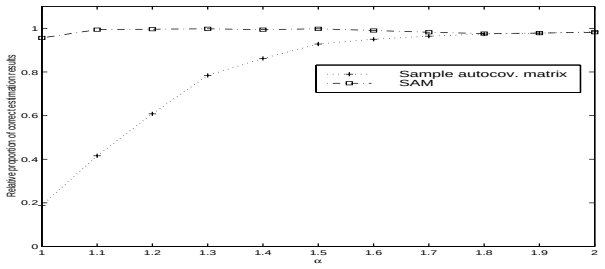

Figure 3: Number of complex exponentials estimation using MDL-criterion.

Proof. See [4].

We can now proceed with the proof. Assume first that the initial phases are deterministic. The distribution of the noise vector is complex Gaussian with the covariance matrix $\sigma^2 I$ and therefore belongs to the family of complex spherically symmetric distributions. To prove the first part, it is therefore enough to prove that the vectors $D(1)\mathbf{a}, \ldots, D(p)\mathbf{a}$ span a p-dimensional subspace i.e. the matrix $W = [D(1)\mathbf{a}, \ldots D(p)\mathbf{a}]$ is of rank p. This is done by writing

$$W = \begin{bmatrix} A_1 e^{j\omega_1} & & 0 \\ & \ddots & \\ 0 & & A_p e^{j\omega_p} \end{bmatrix} \begin{bmatrix} 1 & \cdots & e^{j(p-1)\omega_1} \\ 1 & \cdots & e^{j(p-1)\omega_2} \\ \vdots & \ddots & \vdots \\ 1 & \cdots & e^{j(p-1)\omega_p} \end{bmatrix}$$

$$= FL.$$

Because L^T is a $p \times p$ Vandermonde matrix it follows that L is of rank p. By noting that F is of full rank we get that rank of W is p. Lemma 1 now implies that the $M - p$ smallest eigenvalues of $E\{R_S\}$ are equal and the corresponding eigenvectors are orthogonal to the columns of the matrix B. The second part of the Theorem follows from Theorem 1.8.E in [5].

The proof is completed by noting that if the initial phases are stochastic, the results are true for every realization of $\{\phi_i\}_{i=1}^{p}$, and therefore hold generally.