THE RAO TEST FOR TESTING BANDEDNESS OF COMPLEX-VALUED COVARIANCE MATRIX

Zhenghan Zhu, Steven M. Kay

University of Rhode Island Dept. of Electrical, Computer and Biomedical Engineering 4 East Alumni Ave. Kingston, RI, 02881, USA

ABSTRACT

Banding the inverse of covariance matrix has become a popular technique to estimate a high dimensional covariance matrix from limited number of samples. However, little work has been done in providing a criterion to determine when a matrix is bandable. In this paper, we present a detector to test the bandedness of a Cholesky factor matrix. The test statistic is formed based on the Rao test, which does not require the maximum likelihood estimates under the alternative hypothesis. In many fields, such as radar signal processing, the covariance matrix and its unknown parameters are often complex-valued. We focus on dealing with complex-valued cases by utilizing the complex parameter Rao test, instead of the traditional real Rao test. This leads to a more intuitive and efficient test statistic. Examples and computer simulations are given to investigate the derived detector performance.

Index Terms- Covariance Matrix, Banded, Rao Test

1. INTRODUCTION

In statistical signal processing, such as used in a radar signal processing system, the sample covariance matrix plays an essential role. [1]. It is often estimated from N adjacent sample data vectors $[\mathbf{x}_0 \ \mathbf{x}_1 \ \cdots , \mathbf{x}_{N-1}]$, where \mathbf{x}_n 's are assumed to be $L \times 1$ identical and independent distributed (IID) complexvalued data vectors, with the general maximum likelihood covariance matrix estimate $\hat{\mathbf{C}} = \frac{1}{N} \sum_{n=0}^{N-1} \mathbf{x}_n \mathbf{x}_n^H$ [3], where H denotes hermitian. A good covariance matrix estimate usually requires N to be large. For example, it requires $N \ge 2L$ in space-time adaptive processing (STAP) to have a good clutter covariance matrix estimate [2]. In practice, however, this is not valid due to the nonstationary environment. For example, the data for a STAP system is often nonstationary due to the heterogeneous clutter [1]. The number of data sufficiently IID (homogeneous) can be relatively small $N \le L$ [2].

A popular solution to the problem is adopting banding/tapering techniques. Wu et al. proposed to estimate the covariance matrix by banding the cholesky factor matrix and applying kernel smoothing estimation [4]. Bickel demonstrated that within the bandable class of covariance matrices, the estimator $\hat{\mathbf{C}}^{-1}$ obtained by banding the cholesky factor matrix of the covariance matrix's inverse is consistent [5]. However, little work is available to provide a guideline/criterion on deciding if a covariance matrix or the cholesky factor matrix of its inverse is bandable. Such a criterion is important and useful to decide if the banding technique is a suitable strategy. Other covariance estimation methods, such as modeling the covariance matrix as a timevarying autoregressive moving average (ARMA) model [8] also requires testing to decide if the model is a good fit. Some recent hypothesis tests for bandedness can be found in [6].

In this paper, a new test based on the Rao test is presented to test the bandedness of a Cholesky factor matrix. The Rao test has an asymptotic optimality property for large data records, yet it requires noticeably lower computation cost than some other detectors, ie., generalized likelihood ratio test (GLRT), as it only needs the maximum likelihood estimates (MLE) under the null hypotheses [9]. This property in computational cost of the Rao test can be an advantage in high-dimensional multivariate signal processing. We consider a complex-valued covariance matrix and unknown parameters in this paper. We adopt the complex parameter Rao test, which offers a more intuitive detector than the traditional real Rao test for testing complex-valued parameters [7]. It should be pointed out, however, that the concept of utilizing the Rao test for testing the bandedness of a matrix can also be easily applied to the real-valued covariance matrix case via the real Rao test.

The paper is organized as follows: Section 2 formulates the problem; Section 3 derives the Rao test detector for testing the bandedness of the cholesky factor matrix; Examples and computer simulations for evaluating the detector's performance are given in Section 4; Finally, conclusions are drawn in Section 5.

2. PROBLEM FORMULATION

Assume that we have N IID observed data vectors, $\mathbf{X} = [\mathbf{x}_0^T \ \mathbf{x}_1^T \ \cdots \ \mathbf{x}_{N-1}^T]^T$, where T denotes transpose and each

 \mathbf{x}_n is an $L \times 1$ complex-valued data vector, which obeys a zero-mean multivariate complex Gaussian distribution $\mathbf{x}_n \sim \mathcal{CN}(\mathbf{0}, \mathbf{C})$ for $n = 0, 1, \dots, N-1$, and the \mathbf{x}_n 's are mutually independent. We assume the $N \leq L$ limitation. The $L \times L$ covariance matrices \mathbf{C} is a Hermitian matrix, so its inverse can be decomposed via the Cholesky decomposition as

$$\mathbf{C}^{-1} = \mathbf{D}^H \mathbf{D}$$

where **D** is a lower triangular $L \times L$ matrix with a testing model as follows.

$$\mathbf{D} = \mathbf{D}_B + \sum_{k=1}^M b_k \mathbf{\Phi}_k$$

 \mathbf{D}_B is a known banded lower triangular matrix, with the bandwidth to be m, the b_k 's are unknown complex-valued parameters, and the Φ_k 's are known basis matrices. Specifically,

$$b_{1} = [\mathbf{D}]_{m+2,1}, \qquad \mathbf{\Phi}_{1} = \mathbf{e}_{m+2}\mathbf{e}_{1}^{T}$$

$$b_{2} = [\mathbf{D}]_{m+3,2}, \qquad \mathbf{\Phi}_{2} = \mathbf{e}_{m+3}\mathbf{e}_{2}^{T}$$

$$\vdots \qquad \vdots$$

$$b_{L-m-1} = [\mathbf{D}]_{L,L-m-1}, \qquad \mathbf{\Phi}_{L-m-1} = \mathbf{e}_{L}\mathbf{e}_{L-m-1}^{T}$$

$$b_{L-m} = [\mathbf{D}]_{m+3,1}, \qquad \mathbf{\Phi}_{L-m} = \mathbf{e}_{m+3}\mathbf{e}_{1}^{T}$$

$$b_{L-m+1} = [\mathbf{D}]_{m+4,2}, \qquad \mathbf{\Phi}_{L-m+1} = \mathbf{e}_{m+4}\mathbf{e}_{2}^{T}$$

$$\vdots \qquad \vdots$$

$$b_{M} = [\mathbf{D}]_{L,1}, \qquad \mathbf{\Phi}_{M} = \mathbf{e}_{L}\mathbf{e}_{1}^{T}$$

where $M = \frac{(L-m-1)(L-m)}{2}$ and each \mathbf{e}_k is an $L \times 1$ vector with k^{th} element being one and the rest being all zeros. The objective is to test if the lower triangular Cholesky factor matrix \mathbf{D} is equal to the banded lower triangular matrix \mathbf{D}_B . Let $\mathbf{b} = [b_1 \ b_2 \ \dots \ b_M]^T$. The detection problem is equivalent to choosing between the following hypotheses:

$$\begin{aligned} \mathcal{H}_0 : \mathbf{b} &= \mathbf{0}; \\ \mathcal{H}_1 : \mathbf{b} &\neq \mathbf{0}; \end{aligned}$$

3. THE RAO TEST FOR TESTING THE BANDEDNESS

In this section, we derive the complex parameter Rao test for the aforementioned detection problem. The Rao test attains the asymptotic (as $N \to \infty$) performance as the GLRT but avoids requiring MLEs under the alternative hypothesis \mathcal{H}_1 , so its computation cost is often substantially less than the GLRT. This can be a desirable property in high-dimensional signal processing, such as real-time STAP. The derivation of the Rao test statistics follows. Let $\mathbf{b}^* = [b_1^* \ b_2^* \ \dots \ b_M^*]^T$, where * denotes conjugate, and $\underline{\mathbf{b}} = [\mathbf{b}^T \ \mathbf{b}^H]^T$, which is an $2M \times 1$ complex-valued parameter vector. The complex parameter Rao test detector can be formed [7]

$$T_{R}(\mathbf{X}) = \left. \frac{\partial \ln p(\mathbf{X}; \underline{\mathbf{b}})}{\partial \underline{\mathbf{b}}^{*}} \right|_{\mathbf{b}=\mathbf{0}}^{H} \mathbf{I}^{-1}(\underline{\mathbf{b}}) \Big|_{\mathbf{b}=\mathbf{0}} \left. \frac{\partial \ln p(\mathbf{X}; \underline{\mathbf{b}})}{\partial \underline{\mathbf{b}}^{*}} \right|_{\substack{\mathbf{b}=\mathbf{0} \\ (1)}}$$

where,

$$\frac{\partial \ln p(\mathbf{X}; \underline{\mathbf{b}})}{\partial \underline{\mathbf{b}}} = \left[\frac{\partial \ln p(\mathbf{X}; \underline{\mathbf{b}})}{\partial \mathbf{b}}^T \frac{\partial \ln p(\mathbf{X}; \underline{\mathbf{b}})}{\partial \mathbf{b}^*}^T \right]^T,$$
$$\frac{\partial \ln p(\mathbf{X}; \underline{\mathbf{b}})}{\partial \mathbf{b}} = \left[\frac{\partial \ln p(\mathbf{X}; \underline{\mathbf{b}})}{\partial b_1} \quad \frac{\partial \ln p(\mathbf{X}; \underline{\mathbf{b}})}{\partial b_2} \quad \cdots \quad \frac{\partial \ln p(\mathbf{X}; \underline{\mathbf{b}})}{\partial b_M} \right]^T,$$
$$\frac{\partial \ln p(\mathbf{X}; \underline{\mathbf{b}})}{\partial \mathbf{b}^*} = \left[\frac{\partial \ln p(\mathbf{X}; \underline{\mathbf{b}})}{\partial b_1^*} \quad \frac{\partial \ln p(\mathbf{X}; \underline{\mathbf{b}})}{\partial b_2^*} \quad \cdots \quad \frac{\partial \ln p(\mathbf{X}; \underline{\mathbf{b}})}{\partial b_M^*} \right]^T,$$

are based on Wirtinger derivatives. We next find each element $\frac{\partial \ln p(\mathbf{X};\mathbf{b})}{\partial b_k}$ as follows. Firstly,

$$\ln p(\mathbf{X}; \underline{\mathbf{b}}) = \ln \prod_{n=0}^{N-1} p(\mathbf{x}_n; \underline{\mathbf{b}})$$
$$= \ln \left[\frac{1}{\pi^{NL} \prod_{n=0}^{N-1} \det(\mathbf{C})} \exp(-\sum_{n=0}^{N-1} \mathbf{x}_n^H \mathbf{C}^{-1} \mathbf{x}_n) \right]$$
$$= \ln(\frac{1}{\pi^{NL}}) - \sum_{n=0}^{N-1} \mathbf{x}_n^H \mathbf{D}^H \mathbf{D} \mathbf{x}_n + N \ln \det(\mathbf{D}^H \mathbf{D}),$$
(2)

and

$$\frac{\partial \ln p(\mathbf{X}; \underline{\mathbf{b}})}{\partial b_k} = N \frac{\partial \ln \det(\mathbf{D}^H \mathbf{D})}{\partial b_k} - \sum_{n=0}^{N-1} \frac{\partial \mathbf{x}_n^H \mathbf{D}^H \mathbf{D} \mathbf{x}_n}{\partial b_k}$$
$$= N \frac{\partial \ln \det(\mathbf{D}^H \mathbf{D})}{\partial b_k} - \sum_{n=0}^{N-1} \frac{\partial \mathrm{tr}(\mathbf{D} \mathbf{x}_n \mathbf{x}_n^H \mathbf{D}^H)}{\partial b_k},$$
(3)

for $k = 1, 2, \cdots, M$, where

$$\frac{\partial \ln \det(\mathbf{D}^H \mathbf{D})}{\partial b_k} = \operatorname{tr}(\mathbf{D}^{-1} \mathbf{\Phi}_k), \tag{4}$$

and

$$\frac{\partial \operatorname{tr}(\mathbf{D}\mathbf{x}_{n}\mathbf{x}_{n}^{H}\mathbf{D}^{H})}{\partial b_{k}} = \operatorname{tr}(\mathbf{x}_{n}\mathbf{x}_{n}^{H}\mathbf{D}^{H}\mathbf{\Phi}_{k}).$$
(5)

Thus,

$$\frac{\partial \ln p(\mathbf{X}; \underline{\mathbf{b}})}{\partial b_k} = N \operatorname{tr}(\mathbf{D}^{-1} \boldsymbol{\Phi}_k) - \sum_{n=0}^{N-1} \operatorname{tr}(\mathbf{x}_n \mathbf{x}_n^H \mathbf{D}^H \boldsymbol{\Phi}_k),$$
(6)

Under \mathcal{H}_0 , where $\mathbf{b} = \mathbf{0}$,

$$\frac{\partial \ln p(\mathbf{X}; \underline{\mathbf{b}})}{\partial b_k} \bigg|_{\underline{\mathbf{b}} = \mathbf{0}} = N \operatorname{tr}(\mathbf{D}_B^{-1} \mathbf{\Phi}_k) - \sum_{n=0}^{N-1} \operatorname{tr}(\mathbf{\Phi}_k \mathbf{x}_n \mathbf{x}_n^H \mathbf{D}_B^H)$$
(7)

Also, we have

$$\frac{\partial \ln p(\mathbf{X}; \underline{\mathbf{b}})}{\partial b_k^*} = N \operatorname{tr}(\mathbf{D}^{-H} \boldsymbol{\Phi}_k^H) - \sum_{n=0}^{N-1} \operatorname{tr}(\mathbf{D} \mathbf{x}_n \mathbf{x}_n^H \boldsymbol{\Phi}_k^H),$$
(8)

and its value under \mathcal{H}_0

$$\frac{\partial \ln p(\mathbf{X}; \underline{\mathbf{b}})}{\partial b_k^*} \bigg|_{\mathbf{b}=\mathbf{0}} = N \operatorname{tr}(\mathbf{D}_B^{-H} \mathbf{\Phi}_k^H) - \sum_{n=0}^{N-1} \operatorname{tr}(\mathbf{D}_B \mathbf{x}_n \mathbf{x}_n^H \mathbf{\Phi}_k^H)$$
(9)

We next compute $I(\underline{b})$.

$$\mathbf{I}(\underline{\mathbf{b}}) = E\left(\frac{\partial \ln p(\mathbf{X};\underline{\mathbf{b}})}{\partial \underline{\mathbf{b}}^*} \frac{\partial \ln p(\mathbf{X};\underline{\mathbf{b}})^H}{\partial \underline{\mathbf{b}}^*}\right)$$
$$= \begin{bmatrix} \mathbf{A} & \mathbf{B}^* \\ \mathbf{B} & \mathbf{A}^* \end{bmatrix}$$
$$= \begin{bmatrix} M \times M & M \times M \\ M \times M & M \times M \end{bmatrix}$$
(10)

where,

$$\mathbf{A} = E\left(\frac{\partial \ln p(\mathbf{X}; \mathbf{b})}{\partial \mathbf{b}^*} \frac{\partial \ln p(\mathbf{X}; \mathbf{b})^H}{\partial \mathbf{b}^*}\right)$$
$$\mathbf{B} = E\left(\frac{\partial \ln p(\mathbf{X}; \mathbf{b})}{\partial \mathbf{b}} \frac{\partial \ln p(\mathbf{X}; \mathbf{b})^T}{\partial \mathbf{b}}\right)$$

For each element $[\mathbf{A}]_{k,l}$ and $[\mathbf{B}]_{k,l}$ for $1 \le k, l \le M$, we can compute as follows,

$$\mathbf{A}_{k,l} = -E\left(\frac{\partial^2 \ln p(\mathbf{X}; \underline{\mathbf{b}})}{\partial b_k^* \partial b_l}\right)$$
$$= E\left(\sum_{n=0}^{N-1} \operatorname{tr}(\mathbf{\Phi}_l \mathbf{x}_n \mathbf{x}_n^H \mathbf{\Phi}_k^H)\right)$$
$$= N \operatorname{tr}(\mathbf{\Phi}_l \mathbf{D}^{-1} \mathbf{D}^{-H} \mathbf{\Phi}_k^H)$$
(11)

Under \mathcal{H}_0 , where $\mathbf{b} = \mathbf{0}$, we have

$$\mathbf{A}_{k,l}|_{\mathbf{b}=\mathbf{0}} = N \operatorname{tr}(\mathbf{\Phi}_l \mathbf{D}_B^{-1} \mathbf{D}_B^{-H} \mathbf{\Phi}_k^H)$$
(12)

In a similar fashion, we have

$$\mathbf{B}_{k,l} = -E\left(\frac{\partial^2 \ln p(\mathbf{X}; \underline{\mathbf{b}})}{\partial b_k \partial b_l}\right)$$

= $N \operatorname{tr}(\mathbf{D}^{-1} \mathbf{\Phi}_l \mathbf{D}^{-1} \mathbf{\Phi}_k)$ (13)

and its value under \mathcal{H}_0

$$\mathbf{B}_{k,l}|_{\mathbf{b}=\mathbf{0}} = N \operatorname{tr}(\mathbf{D}_B^{-1} \mathbf{\Phi}_l \mathbf{D}_B^{-1} \mathbf{\Phi}_k)$$
(14)

Using equations (7), (9), (10), (12), (14) and the complex parameter Rao test equation (1) will produce the Rao test statistic.

An explicit example is presented next to evaluate the performance of the detector.

4. NUMERICAL EXAMPLES AND COMPUTER SIMULATIONS

Consider a simple example, where we only have the N = 4observed data set $\mathbf{X} = [\mathbf{x}_0^T \mathbf{x}_1^T \mathbf{x}_2^T \mathbf{x}_3^T]^T$, each \mathbf{x}_n 's is a 4×1 complex-valued IID Gaussian vector, $\mathbf{x}_n \sim C\mathcal{N}(\mathbf{0}, \mathbf{C})$. Also, $\mathbf{C}^{-1} = \mathbf{D}^H \mathbf{D}$, and $\mathbf{D} = \mathbf{D}_B + b_1 \Phi_1$ with $\Phi_1 = \mathbf{e}_4 \mathbf{e}_1^T$ and

$$\mathbf{D}_B = \begin{bmatrix} 0.45 & 0 & 0 & 0 \\ -0.25 + 0.25j & 0.5 & 0 & 0 \\ -0.12 + 0.12j & -0.3 + 0.3j & 0.55 & 0 \\ 0 & -0.15 - 0.15j & 0.2 - 0.2j & 0.6 \end{bmatrix}$$

We are testing if the cholesky factor matrix **D** is banded and equal to the known **D**_B. It is equivalent to testing if $b_1 = 0$ versus $b_1 \neq 0$. The Rao test for this example can be shown to be (15).

To evaluate the Rao test performance for this example, we consider three cases under the alternative hypothesis \mathcal{H}_1 , $b_1 = 0.8 - j$; $b_1 = 0.5 + 0.5j$; $b_1 = -0.2 + 0.4j$ respectively. The receiver operating characteristic (ROC)s, showing the relationship of the probability of detection (P_d) versus the probability of false alarm (P_{fa}) of the derived Rao test is given in Figure 1.



Fig. 1. ROC curve of the Rao test detector with different b_1

The Rao test statistic under the null hypothesis \mathcal{H}_0 is chisquared distributed with one degree of freedom, $T_R(\mathbf{X}) \sim \chi_2^2$. The performance of the Rao test can be found asymptotically or as $N \to \infty$. An estimated probability density function (PDF), shown as a bar plot, and the theoretical PDF $(N \to \infty)$ are shown in Figure 2.

5. CONCLUSIONS

The banding technique have become an important technique in high-dimensional covariance matrix estimation with a limited number of samples. However, before adopting the tech-

$$T_{R}(\mathbf{X}) = \frac{\operatorname{Re}\left\{\left(4\operatorname{tr}(\mathbf{D}_{B}^{-1}\boldsymbol{\Phi}_{1}) - \sum_{n=0}^{3}\operatorname{tr}(\mathbf{x}_{n}\mathbf{x}_{n}^{H}\mathbf{D}_{B}^{H}\boldsymbol{\Phi}_{1})\right)^{2}\operatorname{tr}^{*}(\mathbf{D}_{B}^{-1}\boldsymbol{\Phi}_{1}\mathbf{D}_{B}^{-H}\boldsymbol{\Phi}_{1}^{H})\right\}}{2\left[\left|\operatorname{tr}(\boldsymbol{\Phi}_{1}\mathbf{D}_{B}^{-1}\mathbf{D}_{B}^{-H}\boldsymbol{\Phi}_{1}^{H})\right|^{2} - \left|\operatorname{tr}(\mathbf{D}_{B}^{-1}\boldsymbol{\Phi}_{1}\mathbf{D}_{B}^{-1}\boldsymbol{\Phi}_{1})\right|^{2}\right]} - \frac{\operatorname{Re}\left\{\left(4\operatorname{tr}(\mathbf{D}_{B}^{-1}\boldsymbol{\Phi}_{1}) - \sum_{n=0}^{3}\operatorname{tr}(\mathbf{x}_{n}\mathbf{x}_{n}^{H}\mathbf{D}_{B}^{H}\boldsymbol{\Phi}_{1})\right)^{2}\operatorname{tr}(\mathbf{D}_{B}^{-1}\boldsymbol{\Phi}_{1}\mathbf{D}_{B}^{-1}\boldsymbol{\Phi}_{1})\right\}}{2\left[\left|\operatorname{tr}(\boldsymbol{\Phi}_{1}\mathbf{D}_{B}^{-1}\boldsymbol{\Phi}_{B}^{-H}\boldsymbol{\Phi}_{1}^{H})\right|^{2} - \left|\operatorname{tr}(\mathbf{D}_{B}^{-1}\boldsymbol{\Phi}_{1}\mathbf{D}_{B}^{-1}\boldsymbol{\Phi}_{1})\right|^{2}\right]}$$
(15)



Fig. 2. Estimated and theoretical PDF of the test statistic for the case N = 4

nique, it is important to test if the matrix is "bandable". We have introduced the Rao test of bandedness of Cholesky factor matrix of inverse of the covariance matrix in this paper. The Rao test's computational cost is relatively lower than other detectors such as GLRT, yet with reasonably good performance. A concise form of the Rao test for testing bandedness of a complex-valued covariance matrix with complex-valued unknown parameters is present. An example and a simulation are also given to evaluate the proposed detector. The method can be easily applied to the real-valued covariance matrix and parameters case. Moreover, the detector can be applied to test if any element is zero in a matrix, by changing the basis matrix Φ_k accordingly. The derived detector can be used as a pre-processing stage before adopting banding, or certain modeling method, such as ARMA modeling techniques in covariance matrix estimation.

6. REFERENCES

- Melvin, W.L., "A STAP Overview", *IEEE AES Systems Magazine Special Tutorials Issue*, Vol. 19, No.1, Jan. 2004, pp. 19-35
- [2] Melvin, W. L. and Showman, G. A. "An approach to

knowledge-aided covariance estimation". *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 42, No. 3, July 2006, pp. 1021-1042.

- [3] T. Anderson, An Introduction to Multivariate Statistical Analysis. 3rd ed. Wiley, Hoboken, NJ.
- [4] Wu, W. B. and Pourahmadi, M. "Nonparametric estimation of large covariance matrices of longitudinal data", *Biometrika* 2003, vol.90, pp.831-844.
- [5] Bickel, P. J. and Levina, E. "Regularized estimation of large covariance matrices", *Ann. Statist*.2008, vol.36 pp.199-227.
- [6] Qiu, Y-M and Chen, S. X. "Test for bandedness of high dimensional covariance matrices with bandwidth estimation". *Ann. Stat.* 2012, vol.40, pp.1285-1314.
- [7] S. Kay, Z. Zhu, R.S. Ragahavan, "Complex Parameter Rao Test", submitted to *IEEE Transactions on Signal Processing*.
- [8] A. Wiesel, O. Bibi, and A. Globerson, "Time varying autoregressive moving average models for covariance estimation", *IEEE Transactions on Signal Processing*, vol. 61, no. 11, pp. 2791-2801, 2013
- [9] S. Kay, Fundamentals of Statistical Signal Processing: Detection Theory, Prentice-Hall, Englewood Cliffs, NJ, 1998.