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

Band 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 $[x_0 \ x_1 \ \cdots \ x_{N-1}]$, where $x_n$’s are assumed to be $L \times 1$ identical and independent distributed (IID) complex-valued data vectors, with the general maximum likelihood covariance matrix estimate $\hat{C} = \frac{1}{N} \sum_{n=0}^{N-1} x_n x_n^H$ [3], where $H$ denotes hermitian. A good covariance matrix estimate usually requires $N$ to be large. For example, it requires $N \geq 2L$ 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 $N \leq L$ [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 $\hat{C}^{-1}$ 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 time-varying 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, $X = [x_0^T \ x_1^T \ \cdots \ x_{N-1}^T]^T$, where $T$ denotes transpose and each
x_n is an L \times 1 complex-valued data vector, which obeys a zero-mean multivariate complex Gaussian distribution x_n \sim \mathcal{CN}(0, C) for n = 0, 1, \cdots, N-1, and the x_n's are mutually independent. We assume the N \leq L limitation. The L \times L covariance matrices C is a Hermitian matrix, so its inverse can be decomposed via the Cholesky decomposition as

\[ C^{-1} = D^H D, \]

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

\[ D = D_B + \sum_{k=1}^M b_k \Phi_k \]

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 \Phi_k’s are known basis matrices. Specifically,

\[
\begin{align*}
&b_1 = [D]_{m+2,1}, \quad \Phi_1 = e_{m+2}e_T^1 \\
&b_2 = [D]_{m+3,2}, \quad \Phi_2 = e_{m+3}e_T^2 \\
&\vdots \\
&b_{L-m-1} = [D]_{L-L-m-1}, \quad \Phi_{L-m-1} = e_Le_{T-m-1}^T \\
&b_{L-m} = [D]_{m+3,1}, \quad \Phi_{L-m} = e_{m+3}e_T^1 \\
&b_{L-m+1} = [D]_{m+4,2}, \quad \Phi_{L-m+1} = e_{m+4}e_T^2 \\
&\vdots \\
&b_M = [D]_{L,1}, \quad \Phi_M = e_Le_T^1
\end{align*}
\]

where \( M = \frac{(L-m-1)(L-m)}{2} \) and each \( e_k \) 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 D is equal to the banded lower triangular matrix D_B. Let \( b = [b_1 b_2 \ldots b_M]^T \). The detection problem is equivalent to choosing between the following hypotheses:

\[ H_0 : b = 0; \quad H_1 : b \neq 0; \]

\section{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 \rightarrow \infty) performance as the GLRT but avoids requiring MLEs under the alternative hypothesis \( H_1 \), 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 \( b^* = [b_1^* b_2^* \ldots b_M^*]^T \), where * denotes conjugate, and \( b = [b^T b^H]^T \), which is an 2M \times 1 complex-valued parameter vector. The complex parameter Rao test detector can be formed \cite{7}

\[
T_R(X) = \frac{\partial \ln p(X; b)}{\partial b} \bigg|_{b=0}^H I^{-1}(b) \bigg|_{b=0} \frac{\partial \ln p(X; b)}{\partial b^*} \bigg|_{b=0}\]

where,

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

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

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

\[
\ln p(X; b) = \ln \prod_{n=0}^{N-1} p(x_n; b) = \ln \left( \frac{1}{\pi^NL} \prod_{n=0}^{N-1} \det(C)^{-1} \exp(-\sum_{n=0}^{N-1} x_n^H C^{-1} x_n) \right)
\]

\[
= \ln \left( \frac{1}{\pi^NL} \right) - \sum_{n=0}^{N-1} x_n^H D^H D x_n + N \ln \det(D^H D),
\]

and

\[
\frac{\partial \ln p(X; b)}{\partial b_k} = N \frac{\partial \ln \det(D^H D)}{\partial b_k} = \sum_{n=0}^{N-1} \frac{\partial x_n^H D^H D x_n}{\partial b_k}
\]

\[
= N \frac{\partial \ln \det(D^H D)}{\partial b_k} = \sum_{n=0}^{N-1} \frac{\partial \text{tr}(D x_n x_n^H D^H)}{\partial b_k},
\]

for \( k = 1, 2, \cdots, M \), where

\[
\frac{\partial \ln \det(D^H D)}{\partial b_k} = \text{tr}(D^{-1} \Phi_k),
\]

and

\[
\frac{\partial \text{tr}(D x_n x_n^H D^H)}{\partial b_k} = \text{tr}(x_n x_n^H D^H \Phi_k).
\]

Thus,

\[
\frac{\partial \ln p(X; b)}{\partial b_k} \bigg|_{b=0} = N \text{tr}(D^{-1} \Phi_k) - \sum_{n=0}^{N-1} \text{tr}(x_n x_n^H D^H \Phi_k),
\]

Under \( H_0 \), where \( b = 0 \),

\[
\frac{\partial \ln p(X; b)}{\partial b_k} \bigg|_{b=0} = N \text{tr}(D_B^{-1} \Phi_k) - \sum_{n=0}^{N-1} \text{tr}(\Phi_k x_n x_n^H D_B^H).
\]
Also, we have
\[
\frac{\partial \ln p(X; b)}{\partial b_k} = N \text{tr}(D^{-H} \Phi_k^H) - \sum_{n=0}^{N-1} \text{tr}(DnX_nX_n^H \Phi_k^H),
\]
(8)
and its value under \( H_0 \)
\[
\frac{\partial \ln p(X; b)}{\partial b_k}
= N \text{tr}(D_B^{-H} \Phi_k^H) - \sum_{n=0}^{N-1} \text{tr}(D_BnX_nX_n^H \Phi_k^H)
\]
(9)
We next compute \( I(b) \).
\[
I(b) = E \left( \frac{\partial \ln p(X; b)}{\partial b} \frac{\partial \ln p(X; b)^H}{\partial b^*} \right)
= \begin{bmatrix}
A & B^*
B & A^*
\end{bmatrix}
(10)
\]
where,
\[
A = E \left( \frac{\partial \ln p(X; b)}{\partial b} \frac{\partial \ln p(X; b)^H}{\partial b^*} \right)
\]
\[
B = E \left( \frac{\partial \ln p(X; b)}{\partial b} \frac{\partial \ln p(X; b)^T}{\partial b^*} \right)
\]
For each element \( [A]_{k,l} \) and \( [B]_{k,l} \) for \( 1 \leq k, l \leq M \), we can compute as follows,
\[
A_{k,l} = -E \left( \frac{\partial^2 \ln p(X; b)}{\partial b_k \partial b_l} \right)
= E \left( \sum_{n=0}^{N-1} \text{tr}(\Phi_nX_nX_n^H \Phi_k^H) \right)
= N \text{tr}(\Phi_kD^{-1}D^{-H} \Phi_k^H)
\]
(11)
Under \( H_0 \), where \( b = 0 \), we have
\[
A_{k,l}|_{b=0} = N \text{tr}(\Phi_kD_B^{-1}D_B^{-H} \Phi_k^H)
\]
(12)
In a similar fashion, we have
\[
B_{k,l} = -E \left( \frac{\partial^2 \ln p(X; b)}{\partial b_k \partial b_l} \right)
= N \text{tr}(D^{-1} \Phi_kD^{-1} \Phi_k)
\]
(13)
and its value under \( H_0 \)
\[
B_{k,l}|_{b=0} = N \text{tr}(D_E^{-1} \Phi_kD_B^{-1} \Phi_k)
\]
(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 \( N = 4 \) observed data set \( X = [x_0^T \ x_1^T \ x_2^T \ x_3^T]^T \), each \( x_n \)'s is a \( 4 \times 1 \) complex-valued IID Gaussian vector, \( x_n \sim CN(0, C) \). Also, \( C^{-1} = D^H D \) and \( D = D_B + b_1 \Phi_1 \) with \( \Phi_1 = \text{e}_1 \text{e}_1^* \) and \( D_B \) equal to the known \( B \).

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 \( H_1 \), \( b_1 = 0.8 - j; b_1 = 0.5 + 0.5j; b_1 = -0.2 + 0.4j \) respectively. The receiver operating characteristic (ROC) s, showing the relationship of the probability of detection \( (P_d) \) versus the probability of false alarm \( (P_f) \) of the derived Rao test is given in Figure 1.

![Fig. 1. ROC curve of the Rao test detector with different \( b_1 \)](image)

The Rao test statistic under the null hypothesis \( H_0 \) is chi-squared distributed with one degree of freedom, \( T_R(X) \sim \chi^2_2 \). 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-
The estimated and theoretical PDF of test statistics under $H_0$ with $N=4$

### 6. REFERENCES


