



Testing the Equality Hypothesis on a Cross-Covariance Matrix

Xiaomei Chen, Shibin Zhang

Department of Mathematics, Shanghai Maritime University, Shanghai, China

Email: 1249037246@qq.com, sbzhang@shmtu.edu.cn, changpublis@163.com

How to cite this paper: Chen, X.M. and Zhang, S.B. (2019) Testing the Equality Hypothesis on a Cross-Covariance Matrix. *Open Access Library Journal*, 6: e5584. <https://doi.org/10.4236/oalib.1105584>

Received: July 8, 2019

Accepted: August 9, 2019

Published: August 12, 2019

Copyright © 2019 by author(s) and Open Access Library Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

This paper is concerned about testing whether a cross-covariance matrix deviates from a pre-assigned one or not. For this purpose, a new test statistic is constructed based on the Frobenius norm of the difference between the sample cross-covariance matrix and the pre-assigned matrix. The test is implemented by applying the parametric bootstrap scheme. We conduct a simulation study to examine the performance of the test and compare it with other competitive tests. As multiple simulation examples show, our empirical powers are clearly superior to others in detecting any deviation of the cross-covariance from the pre-assigned matrix. In addition, the proposed test is insensitive to non-cross-covariance elements in the covariance matrix. As an illustration, we also investigate its performance in testing pairwise time-reversibility.

Subject Areas

Mathematical Statistics

Keywords

Cross-Covariance Matrix, Parametric Bootstrap, Spectral Analysis, Time Reversibility

1. Introduction

Tests of covariance matrices in multivariate statistical analysis have wide applications in many fields of research and practice, such as target detection [1], face recognition [2] and so on. They have already attracted considerable interests since the 1940s. However, most of the existing researches on this topic focus on testing for covariance matrix other than cross-covariance matrix. In some circumstances, not all the entries in the covariance matrix are concerned, thus

testing for a cross-covariance matrix being equal to a specified one becomes an important issue. For instance, when testing for time-reversibility (see Section 3 for more details), we can transform the problem into the one of cross-covariance matrix test. Therefore, like the covariance matrix test, it is also of great practical interest to develop methods for the cross-covariance matrix test.

Over the past several years, many types of statistics have been proposed to test various equalities of covariance matrices. The first type is a class of statistics based on the likelihood ratio (LR). Mauchly [3] was one of the earlier attempts whose approach was based on the likelihood ratio. The statistic of Mauchly depends on the determinant and the trace of sample covariance matrix. It requires that the sample covariance matrix is non-singular, which is the case with probability one when the sample size is larger than the dimension. Gupta and Xu [4] generalized the likelihood ratio test to non-normal distributions by deriving the asymptotic expansion of the test statistic under the null hypothesis when the sample size is moderate. Latterly, Jiang *et al.* [5] proved that the likelihood ratio test statistic has an asymptotic normal distribution under two different assumptions by the aid of Selberg integrals. Also, the first type of statistics can be extended to analyze high-dimensional data. For instance, Bai *et al.* [6] used central limit theorems for linear spectral statistics of sample covariance matrices and of random F-matrices, and proposed a modification of the likelihood ratio test to cope with high-dimensional effects. In the following, Niu *et al.* [7] considered testing mean vector and covariance matrix simultaneously with high-dimensional non-Gaussian data. Niu *et al.* applied the central limit theorem for linear spectral statistics of sample covariance matrices and established new modification for the likelihood ratio test. The second type is a class of statistics based on empirical distance. Let Z_1, Z_2, \dots, Z_N be a p -dimensional random sample drawn from a normal distribution with mean vector μ and covariance matrix Σ . Nagao [8] proposed a test statistic

$$T_1 = \frac{n}{2} \text{tr} \left(S \Sigma_0^{-1} / n - I \right)^2, \quad (1)$$

to test the null hypothesis $H_0 : \Sigma = \Sigma_0$ versus the alternative $H_1 : \Sigma \neq \Sigma_0$, where $S = \sum_{i=1}^N (Z_i - \bar{Z})(Z_i - \bar{Z})'$ with $\bar{Z} = N^{-1} \sum_{i=1}^N Z_i$, $n = N - 1$ and I is the identity matrix. Thus the null hypothesis should be rejected when the observed value of T_1 exceeds a pre-assigned level of significance. The third type statistic is based on the largest eigenvalue of the covariance matrix and the random matrix theory. For instance, Cai *et al.* [9] studied the limiting laws of the coherence of an $n \times p$ random matrix in the high-dimensional setting that p can be much larger than n , then Cai *et al.* considered testing the structure of the covariance matrix of a high-dimensional Gaussian distribution, where the random matrix plays a crucial role in the construction of the test. The last type is a statistic based on the examination of a fixed column of the sample covariance matrix. Gupta and Bodnar [10] proposed an exact test on the structure of the covariance matrix. The test statistic of Gupta is based on the examination of a fixed column of the

sample covariance matrix, and it can also be applied if the sample size is much smaller than the dimension.

The above mentioned statistics for covariance matrix test are applicable when the dispersion matrix has the Wishart distribution or the distribution of the test statistic is derivable, thus the asymptotic properties of these statistics can be obtained. In many circumstances, the asymptotic distribution of the test statistic is complicated in the absence of strict normality or when the Wishart distribution is unavailable. In this paper, we provide a new method for testing cross-covariance matrix other than covariance matrix, which can be more efficient in some problems that the variance are not concerned. Moreover, the proposed test is independent of Wishart distribution but can be implemented by parametric bootstrap scheme.

The proposed statistic is based on the Frobenius norm of the difference between the sample cross-covariance matrix and the given matrix. Theoretically, it can detect any deviation of the cross-covariance matrix from a pre-assigned one. Several numerical examples show that it is more powerful in testing a cross-covariance matrix deviating from the pre-assigned matrix than some other competitive methods.

Recently, tests of time-reversibility (TR) have drawn much attention due to that time reversibility is a necessary condition for an independent and identically distributed (i.i.d) sequence. As is known, i.i.d sequence and stationary Gaussian models are time reversible. Otherwise, a linear, non-Gaussian process is time-irreversible, except when its coefficients satisfy particular constraints [11]. Several tests for TR have been proposed to be applied as tests for specification check in model construction [12] [13] [14] [15]. In this paper, the TR test method is based on the copula spectral density kernel (CSDK) proposed by Dette *et al.* [16], which is more informative than the traditional spectral density, the CSDK captures serial dependence more than covariance-related. The CSDK $f_{q_{\tau_1}, q_{\tau_2}}(\omega)$ (defined in (17)) is indexed by couple (q_{τ_1}, q_{τ_2}) of quantile levels, where $(\tau_1, \tau_2) \in [0, 1]^2$ and $q_{\tau_i} = F^{-1}(\tau_i)$ ($i=1, 2$), with $F(x)$ being the one-dimensional marginal cumulative distribution function of a strictly stationary univariate process $\{x_t\}_{t \in \mathbb{Z}}$. Obviously, the time series is pairwise time reversible if and only if $\Im f_{q_{\tau_1}, q_{\tau_2}}(\omega) = 0$ for all ω and all $(\tau_1, \tau_2) \in [0, 1]^2$, where $\Im a$ is the imaginary part of a complex number a . Thus the imaginary part of CSDK is equal to zero if $\{x_t\}_{t \in \mathbb{Z}}$ is time reversible. So we can transform the problem of testing pairwise time-reversibility into one of testing the imaginary part of CSDK being zero. By Theorem 3.3 of Dette *et al.* [16], we derived a covariance matrix $\Sigma_2(\omega)$ (defined in (24)), we find that time-reversibility indicates the cross-covariance matrix in $\Sigma_2(\omega)$ is equal to a zero matrix. Theoretically, we can transform the problem of testing pairwise time-reversibility into that of testing for the specification of a cross-covariance matrix.

Throughout the paper, we denote by $\stackrel{d}{=}$ equality in distribution, and define $\Re a$ and $\Im a$ as the real part and imaginary part of a complex number a , respectively. For matrix notation, I_m and O_m denote the $m \times m$ identity ma-

trix and $m \times m$ zero matrix, respectively; $\det(M)$ and $\text{tr}(M)$ represent the determinant and trace of the matrix M , respectively; $\|M\|_F$ indicates the Frobenius norm of M . $\chi^2(q)$ and t_q denote a chi-square distribution and a Student t distribution with q degrees of freedom, respectively.

The rest of the paper is organized as follows. Section 2 presents the test statistic with the bootstrap scheme in computing the p -value of the cross-covariance matrix test. Section 3 reports empirical results for examining performance of the proposed test by using simulated data. Section 4 illustrates the applications in detecting any deviation from time-reversibility of a time series. Section 5 contains our conclusions.

2. Test Statistic and Its Distributional Approximation

2.1. Test Statistic

Let Z_1, \dots, Z_N be an independent sample from $\mathcal{N}_p(\boldsymbol{\mu}, \Sigma)$, $p \geq 2$, an multivariate normal distribution with mean vector $\boldsymbol{\mu}$, covariance matrix Σ , where Z_i is expressed by $\begin{pmatrix} X_i \\ Y_i \end{pmatrix}$, $i = 1, \dots, N$, and Σ is a blocked matrix given by

$$\Sigma = \begin{pmatrix} \Sigma_x & \Gamma \\ \Gamma' & \Sigma_y \end{pmatrix}, \quad (2)$$

with Γ being the cross-covariance matrix. In this section, we consider the problem of testing

$$H_0: \Gamma = \Gamma_0 \text{ versus } H_1: \Gamma \neq \Gamma_0.$$

The test statistic is constructed based on the Frobenius norm of the difference between the sample cross-covariance matrix and the given matrix. In the derivation, no assumption on p , like $n \geq p$ or $n < p$, is required. Since the Wishart distribution is not achieved here, we implement the derivation by the aid of parametric bootstrap scheme. We define the test statistic

$$T = N \left\| \frac{1}{N} S_{xy} - \Gamma_0 \right\|_F^2, \quad (3)$$

where

$$S_{xy} = \sum_{i=1}^N (X_i - \bar{X}) \cdot (Y_i - \bar{Y})', \quad (4)$$

with $\bar{X} = N^{-1} \sum_{i=1}^N X_i$ and $\bar{Y} = N^{-1} \sum_{i=1}^N Y_i$. In (3),

$\left\| \frac{1}{N} S_{xy} - \Gamma_0 \right\|_F^2 = \text{tr} \left[\left(\frac{1}{N} S_{xy} - \Gamma_0 \right)^T \left(\frac{1}{N} S_{xy} - \Gamma_0 \right) \right]$, which can detect any deviation of cross-covariance from the pre-specified matrix Γ_0 .

2.2. Bootstrap Approximation of the Null Distribution

Let Z_1, Z_2, \dots, Z_N be an independent sample that are drawn from $\mathcal{N}_p(\boldsymbol{\mu}, \Sigma)$. $\hat{\Sigma}_x$ and $\hat{\Sigma}_y$ are the estimator of the parameters Σ_x and Σ_y . Suppose that the

pseudo data set Z_1^*, \dots, Z_N^* was resample from $\mathcal{N}_p(\boldsymbol{\mu}, \Sigma^*)$, where

$$\Sigma^* = \begin{pmatrix} \hat{\Sigma}_x & \Gamma \\ \Gamma' & \hat{\Sigma}_y \end{pmatrix}, \quad (5)$$

and Z_i^* is expressed by $\begin{pmatrix} X_i^* \\ Y_i^* \end{pmatrix}, i=1, \dots, N$. The bootstrap statistic is defined as

$$T^* = N \left\| \frac{1}{N} S_{xy}^* - \Gamma_0 \right\|_F^2, \quad (6)$$

where

$$S_{xy}^* = \sum_{i=1}^N (X_i^* - \bar{X}^*) \cdot (Y_i^* - \bar{Y}^*)', \quad (7)$$

with $\bar{X}^* = N^{-1} \sum_{i=1}^N X_i^*$ and $\bar{Y}^* = N^{-1} \sum_{i=1}^N Y_i^*$.

To study the bootstrap approximation of the null distribution, we need the following conditions:

Condition (A1) Let \tilde{F} be the cumulative distribution function (cdf) of the bootstrap statistic T^* , F_M be the empirical distribution function of the bootstrap sample T_1^*, \dots, T_M^* . Let

$$W_{\tilde{F}M}(t) = \sqrt{M} \{F_M(t) - \tilde{F}(t)\}. \quad (8)$$

We assume the value of the equality (8) vanishes as t goes to infinity.

Condition (A2) As M tends to infinity, $W_{\tilde{F}M}$ converges weakly to $B(F)$ in distribution, provided that $\sup_{-\infty < x < +\infty} |\tilde{F}(x) - F(x)| \rightarrow 0$ a.s., where F denotes the cdf of the statistic T and B is the Brownian bridge on $[0, 1]$.

Theorem 1. Suppose F is nondegenerate. Suppose also that $0 < \alpha < 1$ be the nominal size of the test. Under conditions (A1) and (A2), if $c(\tilde{F})$ satisfies

$$P\{M^{1/2} \sup_x |F_M(x) - \tilde{F}(x)| \leq c(\tilde{F})\} \rightarrow 1 - \alpha, \quad (9)$$

then

$$P\{M^{1/2} \sup_x |\tilde{F}(x) - F(x)| \leq c(\tilde{F})\} \rightarrow 1 - \alpha. \quad (10)$$

The result (9) is almost immediate from Corollary 4.2 of Bickel *et al.* [17]. By Lemma 8.11 of Bickel *et al.* [17], we obtain that $\sup_x |B(F(x))|$ has a continuous distribution, $c(\tilde{F})$ converges to the $(1-\alpha)$ -quantile of the law of $\sup_x |B(F(x))|$.

2.3. Algorithm for Calculating Test p -Value

In order to carry out the parametric bootstrap procedure for the proposed test, we present the following simulation steps. The bootstrap p -value is approximated by the following procedure.

Step 1. Calculate an observation T_{obs} of statistic T .

Step 2. Estimate the covariance matrix Σ_x and Σ_y by sample covariance matrices, say $\hat{\Sigma}_x$ and $\hat{\Sigma}_y$.

Step 3. Resample from $\mathcal{N}_p(\boldsymbol{\mu}, \Sigma^*)$ and calculate the value of bootstrap sta-

tistic T^* , where Σ^* is defined in (5).

Step 4. Repeat Step 3 M times, and compute the p -value by $p\text{-value} = \frac{1}{M} \sum_{i=1}^M \mathbb{I}_{\{T_i^* > T_{obs}\}}$, where \mathbb{I}_A denotes an indicator on the set A , which equals 1 when A occurs, and 0 otherwise.

3. A Simulation Study

3.1. Comparison Study

We briefly describe the tests which are compared in the current paper, the modified LR test [18], and the test of Nagao [8]. These two tests are used to test

$$H_0: \Sigma = \Sigma_0 \text{ versus } H_1: \Sigma \neq \Sigma_0.$$

Let Z_1, \dots, Z_N be an independent sample from $\mathcal{N}_p(\mu, \Sigma)$, the modified LR test statistic is based on

$$\lambda_1^* = e^{\frac{1}{2}pn} \left(\det(S\Sigma_0) e^{-\text{tr}(S\Sigma_0^{-1})} \right)^{\frac{1}{2^n}}, \quad (11)$$

where $S = (1/n)A$, $n = N - 1$, $A = \sum_{i=1}^N (Z_i - \bar{Z})(Z_i - \bar{Z})'$, and $\bar{Z} = N^{-1} \sum_{i=1}^N Z_i$. As is known, $A \sim W_p(N-1, \Sigma)$, where the $W_p(N-1, \Sigma)$ stands for p -dimension Wishart distribution with $N-1$ degrees of freedom and covariance matrix Σ . Anderson [19] derived the limiting distribution of the modified LR test statistic $-2 \log \lambda_1^*$ with the help of Wishart distribution, when $\Sigma = I_p$, $-2 \log \lambda_1^*$ is asymptotically distributed as $\chi^2\left(\frac{1}{2}p(p+1)\right)$, where $\log a$ denotes the logarithmic function based on natural logarithm e . Then, Nagao [8] proposed a test statistic T_1 (defined in (1)) which can be regarded as a measure of departure from the null hypothesis.

In what follows, we propose the statistic $T(3)$ to test

$$H_0: \Gamma = O_{\frac{p}{2}} \text{ versus } H_1: \Gamma \neq O_{\frac{p}{2}}. \quad (12)$$

So far, no test methods are available for this problem. Thus, we choose statistics $-2 \log \lambda_1^*$ and T_1 to test

$$H_0: \Sigma = I_p \text{ versus } H_1: \Sigma \neq I_p, \quad (13)$$

for the comparative study, where the cross-covariance matrix in Σ is equal to $O_{\frac{p}{2}}$. Testing the structure of the covariance matrix can also detect the deviation of cross-covariance from the pre-specified matrix. For a pre-specified level of significance α ($0 < \alpha < 1$), the null hypothesis in (13) is rejected if

$$-2 \log \lambda_1^* > \chi_{1-\alpha}^2 \left(\frac{1}{2}p(p+1) \right), \quad (14)$$

or

$$T_1 > T_{1-\alpha}(n), \quad (15)$$

where $T_{1-\alpha}(n)$ denotes the $1-\alpha$ quantile of the empirical distribution of statistic T_1 .

We employ simulation data to evaluate the performance of the proposed statistic T , statistics $-2\log\lambda_1^*$ and T_1 when applied to test the hypotheses (12) and (13) at a significant level of $\alpha = 0.05$. Empirical sizes and powers of the proposed test are computed based on $M = 250$ resample times and 500 repetition times, and that of tests (14) and (15) are based on 500 Monte Carlo replications. In the simulation study we choose mean vector $\mu = O_p$, dimension $p = 4$, and take sample size $n \in \{50, 100, 200, 400, 800\}$. The results are shown in **Table 1**.

Results in **Table 1** show that each empirical type I error rate of above three different statistics is very close to the pre-specified nominal value. Also, the proposed test has higher empirical power than its counterpart tests (14) and (15). With increasing sample size, the change in empirical power of test (15) is barely noticeable while the performance of the proposed test improves significantly, which means statistic T_1 is not sensitive to the change of the cross elements of the covariance matrix while statistic T can detect any deviation of cross-covariance from the pre-specified matrix Γ_0 . Thus, when the variance in the covariance matrix is not concerned, we recommend that statistic T can be applied to testing the equality hypothesis about a cross-covariance matrix.

3.2. Bootstrap Asymptotic Study

In this section, we employ simulation data to investigate whether the performance of proposed test is sensitive to the block matrices in the diagonal. For this purpose, we consider two choices of Σ : one is

$$\Sigma_1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

and the other is

$$\Sigma_2 = \begin{pmatrix} 1 & 0.5 & 0 & 0 \\ 0.5 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0.5 \\ 0 & 0 & 0.5 & 1 \end{pmatrix}.$$

We propose statistic T to test

$$H_0: \Gamma = O_2 \text{ versus } H_1: \Gamma \neq O_2. \quad (16)$$

For the two covariance matrices mentioned above, we run a simulation with $M = 250$ resample times and 500 repetition times to obtain the empirical sizes and powers of the proposed test at significant level $\alpha = 0.01, 0.05, 0.1$, where we take sample size $n \in \{64, 128, 256, 512, 1024\}$ and choose mean vector $\mu = O_p$. The results are shown in **Table 2** and **Table 3**.

For each sample size $n = 128, 256, 512, 1024$ and each nominal size $\alpha = 0.01, 0.05, 0.1$, **Table 2** shows the empirical rejection probabilities of the proposed test. We present simulation results of Σ_1 and Σ_2 in the first panel

Table 1. Rejection probabilities of the proposed test, tests (14) and (15) from simulated data.

n	Statistic $-2\log \lambda_1^*$		Statistic T_1		Statistic T'	
	Size	Power	Size	Power	Size	Power
50	0.052	0.230	0.059	0.062	0.042	0.286
100	0.053	0.420	0.048	0.050	0.046	0.584
200	0.042	0.782	0.042	0.044	0.048	0.906
400	0.046	0.989	0.056	0.058	0.058	0.996
800	0.054	1	0.054	0.052	0.046	1

Table 2. Probability of committing the type I error of the proposed test in testing (16) for two different Σ .

Σ	α	n			
		128	256	512	1024
$\Sigma = \Sigma_1$	$\alpha = 0.01$	0.008	0.012	0.014	0.012
	$\alpha = 0.05$	0.056	0.052	0.054	0.050
	$\alpha = 0.1$	0.112	0.090	0.102	0.088
$\Sigma = \Sigma_2$	$\alpha = 0.01$	0.010	0.012	0.012	0.010
	$\alpha = 0.05$	0.044	0.058	0.050	0.046
	$\alpha = 0.1$	0.092	0.114	0.092	0.088

Table 3. Empirical rejection probabilities of the proposed test in testing (16) for two different Σ .

Σ	$n = 64$	$n = 128$	$n = 256$	$n = 512$	$n = 1024$
$\Sigma = \Sigma_1$	0.402	0.734	0.962	1	1
$\Sigma = \Sigma_2$	0.286	0.594	0.954	1	1

and the second panel, respectively. We see that each empirical type I error rate of two different cases is close to their nominal sizes. For the nominal size $\alpha = 0.05$, **Table 3** shows how the empirical rejection probability of the proposed test changes with respect to five different sample sizes $n = 64, 128, 256, 512, 1024$. Although the block matrices in the diagonal are different, their empirical powers improve significantly with increasing sample size n . Last but not least, when $n = 512$, both the empirical powers of above two cases reach the maximum. By our simulation experiments in **Table 2** and **Table 3**, we find that the proposed test is not sensitive to the block matrices in the diagonal. It is due to that the proposed statistic T' depends only on the sample cross-covariance matrix. Thus we can conclude that our proposed test still achieve good performance though the change of the variance components in the covariance matrix take place.

4. An Empirical Application: Testing for Pairwise Time-Reversibility

4.1. Time Reversible Time Series and Prior Specification

A formal statistical definition of pairwise time-reversibility is defined as follows.

Definition 1. A time series $\{x_t\}_{t \in \mathbb{Z}}$ is pairwise time reversible if for all positive integers k , the random vectors $(x_t, x_{t+k})^T$ and $(x_t, x_{t-k})^T$ have the same joint probability distributions.

Under this definition, one can show that pairwise time reversibility implies stationarity. Likewise, nonstationarity implies time irreversibility [14]. Clearly, $\{x_t\}$ is time reversible when $\{x_t\}$ is i.i.d. Thus, for the study of testing time-reversibility, the pairwise case is generally considered. For instance, Ramsey *et al.* [12] proposed a pairwise TR test statistic consists of a sample estimate of the symmetric-bicovariance function given by the difference between two bicovariances of (x_t, x_{t-k}) . Later, Chen *et al.* [11] considered the pairwise time-reversibility and proposed a new test aiming at the symmetrical distribution of $(x_t - x_{t-k})$ rather than moments. In the following, Dette *et al.* [16] briefly analyzed the pairwise time-reversibility of four different time series models by the aid of quantile-based spectral analysis.

In this section, we primarily focus on testing for pairwise time-reversibility and the test method is based on *copula spectral density kernel* (CSDK) proposed by Dette *et al.* [16]. Let $\{x_t\}_{t \in \mathbb{Z}}$ be a strictly stationary univariate process, the CSDK is defined as

$$f_{q_{\tau_1}, q_{\tau_2}}(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \gamma_k^U(\tau_1, \tau_2) e^{-ik\omega}, \quad (\tau_1, \tau_2) \in (0, 1)^2, \quad (17)$$

where q_{τ_i} is the τ_i -quantile of the marginal distribution of the process $\{x_t\}_{t \in \mathbb{Z}}$, i.e. $q_{\tau_i} = F^{-1}(\tau_i)$, $i = 1, 2$. The $\gamma_k^U(\tau_1, \tau_2)$ (*copula cross-covariance kernel*) of lag $k \in \mathbb{Z}$ is also introduced by Dette *et al.* [16] which is defined as

$$\gamma_k^U(\tau_1, \tau_2) := \text{Cov}\left(\mathbb{I}_{(0, \tau_1]}(U_t), \mathbb{I}_{(0, \tau_2]}(U_{t-k})\right), \quad (18)$$

where $U_t := F(x_t)$, F denotes one-dimensional marginal cumulative distribution function of the process $\{x_t\}_{t \in \mathbb{Z}}$. Compared with traditional covariances, the concept of *copula cross-covariance kernel* is proper for describing a serial copula.

The collection of CSDKs for different (τ_1, τ_2) provides a full characterization of the copulas associated with the pairs (x_t, x_{t-k}) , and accounts for many important dynamic features of $\{x_t\}$, such as changes in the conditional shape (skewness, kurtosis), time-irreversibility, or dependence in the extremes that the traditional second-order spectra cannot capture [20].

4.2. Test for Time-Reversibility

In the sequel, we will concentrate on testing for pairwise time-reversibility. Obviously, we have $(x_t, x_{t+k}) \stackrel{d}{=} (x_t, x_{t-k})$ for all $k \in \mathbb{Z}$ if and only if

$\Im f_{q_{\tau_1}, q_{\tau_2}}(\omega) = 0$ for all ω and all $(\tau_1, \tau_2) \in [0, 1]^2$. For the purpose of pairwise

time-reversibility test, we consider the problem of testing

$$H_0 : \Im f_{q_{\tau_1}, q_{\tau_2}}(\omega) = 0 \text{ versus } H_1 : \Im f_{q_{\tau_1}, q_{\tau_2}}(\omega) \neq 0, \quad (19)$$

for all $\omega \in (0, \pi)$. The method introduced by Dette *et al.* [16] for estimating the $f_{q_{\tau_1}, q_{\tau_2}}(\omega)$ is first to calculate the *rank-based Laplace periodogram* (RLP) and then to smooth it to obtain the consistent estimator. Let $\{x_1, \dots, x_n\}$ be the observation from a strictly stationary univariate process $\{x_t\}$. Like Dette *et al.* [16], we define $\hat{\mathbf{b}}$ by

$$(\hat{a}_{n,R}^\tau(\omega_{j,n}), \hat{\mathbf{b}}_{n,R}^\tau(\omega_{j,n})) := \underset{a, \mathbf{b}^\top \in \mathbb{R}^3}{\operatorname{Argmin}} \sum_{t=1}^n \rho_\tau \left(n^{-1} R_t^{(n)} - (a, \mathbf{b}^\top) \mathbf{c}_t(\omega_{j,n}) \right), \quad (20)$$

where

$$\rho_\tau(x) := x \left(\tau - \mathbb{I}_{(-\infty, 0]}(x) \right), \quad (21)$$

is the so-called check function [21], $\mathbf{c}_t(\omega_{j,n}) := (1, \cos(t\omega_{j,n}), \sin(t\omega_{j,n}))^\top$,

$\omega_{j,n} \in \mathcal{F}_n := \left\{ 2\pi j/n \mid j = 1, \dots, \left\lfloor \frac{n-1}{2} \right\rfloor - 1, \left\lfloor \frac{n-1}{2} \right\rfloor \right\}$. We extend the definition of $\hat{\mathbf{b}}_{n,R}^\tau(\omega_{j,n})$ to a piecewise constant function on $\Omega_n := \left(\omega_{j,n} - \frac{2\pi}{n}, \omega_{j,n} + \frac{2\pi}{n} \right]$ as follows:

$$\hat{\mathbf{b}}_{n,R}^\tau(\omega) = \hat{\mathbf{b}}_{n,R}^\tau(\omega_{j,n}), \text{ if } \omega_{j,n} - \frac{2\pi}{n} < \omega < \omega_{j,n} + \frac{2\pi}{n}. \quad (22)$$

Let $\tau = \{\tau_1, \tau_2\}^\top$. We denote $\hat{\mathbf{b}}_{n,R}^\tau(\omega) = (\hat{\mathbf{b}}_{n,R}^{\tau_1}(\omega)^\top, \hat{\mathbf{b}}_{n,R}^{\tau_2}(\omega)^\top)^\top$, where $\sqrt{n} \hat{\mathbf{b}}_{n,R}^{\tau_1}(\omega) = (\mathbf{Z}_{11}(\omega), \mathbf{Z}_{12}(\omega))^\top$ and $\sqrt{n} \hat{\mathbf{b}}_{n,R}^{\tau_2}(\omega) = (\mathbf{Z}_{21}(\omega), \mathbf{Z}_{22}(\omega))^\top$. Then, under standard mixing conditions (Theorem 3.3 of Dette *et al.* [16]), for $\omega \in (0, \pi)$, $(\mathbf{Z}_{11}(\omega), \mathbf{Z}_{12}(\omega), \mathbf{Z}_{21}(\omega), \mathbf{Z}_{22}(\omega))^\top$ converges to a zero-mean real Gaussian distribution with covariance matrix

$$\Sigma_4(\omega) := 4\pi \begin{pmatrix} f_{q_{\tau_1}, q_{\tau_1}}(\omega) & 0 & \Re f_{q_{\tau_1}, q_{\tau_2}}(\omega) & \Im f_{q_{\tau_1}, q_{\tau_2}}(\omega) \\ 0 & f_{q_{\tau_1}, q_{\tau_1}}(\omega) & -\Im f_{q_{\tau_1}, q_{\tau_2}}(\omega) & \Re f_{q_{\tau_1}, q_{\tau_2}}(\omega) \\ \Re f_{q_{\tau_1}, q_{\tau_2}}(\omega) & -\Im f_{q_{\tau_1}, q_{\tau_2}}(\omega) & f_{q_{\tau_2}, q_{\tau_2}}(\omega) & 0 \\ \Im f_{q_{\tau_1}, q_{\tau_2}}(\omega) & \Re f_{q_{\tau_2}, q_{\tau_2}}(\omega) & 0 & f_{q_{\tau_2}, q_{\tau_2}}(\omega) \end{pmatrix}. \quad (23)$$

Not all the entries in $\Sigma_4(\omega)$ need to be concerned when testing hypotheses (19), so we consider the block covariance matrix $\operatorname{Cov}(\mathbf{Z}_{11}(\omega), \mathbf{Z}_{22}(\omega))$,

$$\Sigma_2(\omega) := 4\pi \begin{pmatrix} f_{q_{\tau_1}, q_{\tau_1}}(\omega) & \Im f_{q_{\tau_1}, q_{\tau_2}}(\omega) \\ \Im f_{q_{\tau_1}, q_{\tau_2}}(\omega) & f_{q_{\tau_2}, q_{\tau_2}}(\omega) \end{pmatrix}. \quad (24)$$

Thus, for the test of hypotheses (19), we can transform the problem into one of testing the cross-covariance matrix in (24) being a one dimensional zero matrix. Here, we define random vectors X and Y as

$$X = \operatorname{sgn}(\hat{f}_{\tau_1, \tau_1}(\omega)) \frac{\mathbf{Z}_{11}(\omega)}{\sqrt{4\pi |\hat{f}_{\tau_1, \tau_1}(\omega)|}}, Y = \operatorname{sgn}(\hat{f}_{\tau_2, \tau_2}(\omega)) \frac{\mathbf{Z}_{22}(\omega)}{\sqrt{4\pi |\hat{f}_{\tau_2, \tau_2}(\omega)|}},$$

where the smoothed rank-based Laplace periodogram $\hat{f}_{n,\tau_1,\tau_2}(\omega)$ denotes a consistent estimator of CSDK, and

$$\text{sgn}(\hat{f}_{n,\tau_i,\tau_i}(\omega)) = \begin{cases} 1, & \text{if } \hat{f}_{n,\tau_i,\tau_i}(\omega) > 0, \\ -1, & \text{if } \hat{f}_{n,\tau_i,\tau_i}(\omega) < 0, \end{cases} \quad i = 1, 2 \quad (25)$$

Let Z_1, \dots, Z_N be an independent sample, where Z_i is expressed by $\begin{pmatrix} X_i \\ Y_i \end{pmatrix}, i = 1, \dots, \omega$. The test p -value is approximated by the previous bootstrap scheme.

4.3. An Illustration Example

Due to lack of the accurate values of CSDKs, it is difficult for us to evaluate our methodology by using the general models by simulation [20]. We consider two AR(1) models (models 1 and 2) with the form

$$x_t = -0.3x_{t-1} + \varepsilon_t, \quad (26)$$

since their CSDKs can be computed numerically. In model 1, $\varepsilon_t, t = 1, 2, \dots$ are independent $\mathcal{N}(0, 1)$ -distributed random variables, while independent Student t -distributed with 1 degree of freedom in model 2. For model 1, the imaginary component of CSDK is vanishing, which reflects that the process is time-reversible; for model 2, there exists a time-irreversible impact of extreme values on the central ones [16].

Recently, Dette *et al.* [16] and Kley *et al.* [22] proposed to estimate the CSDKs by smoothing RLP, which can be defined by quantile regression (QR) or clipped time series (CT). The finite sample performance of the smoothed RLP can be conducted using the *R* package *quantspec* [22]. This makes the method of smoothed (QR- or CT-based) RLP serve a good reference to calculate the test p -value. We take the smoothed QR-based RLP to cope with the pairwise TR test. For each generated pseudo-random time series, we computed the smoothed QR-based RLP using the Epanechnikov kernel and bandwidth $bw = 0.07$. For each generated dataset and each pair of (τ_1, τ_2) , the test p -value is computed by the previous bootstrap procedure.

For each of those two models, we generated 512 and 1024 dataset with each containing pseudo-random time series of lengths $n = 1024$ and $n = 2048$ respectively. We set $\tau = (0.1, 0.5, 0.9)^T$. A boxplot that is drawn is based on 50 realizations of the log scale of the realized p -value. For each pairs $(\tau_1, \tau_2) = (0.1, 0.9), (0.1, 0.5), (0.5, 0.9)$, the boxplots of two models are presented on the left, middle and right, respectively. For each boxplot, the median, extreme points and box shaped by lower and upper quartiles are marked in **Figure 1** and **Figure 2**.

Next we discuss the simulation results in the case of an AR(1) process. We find from **Figure 1** that the lower quartile of boxplots of two models is greater than $\log(\alpha)$ at quantile pair $(0.1, 0.9)$, which means the null hypothesis in (19) cannot be rejected, *i.e.* $P(x_t \leq q_{0.1}, x_{t+k} \leq q_{0.9})$ is approximately equal to

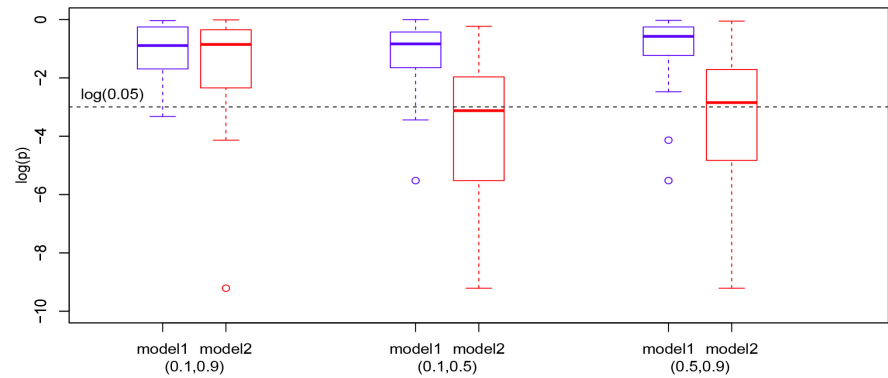


Figure 1. Boxplots of the estimated $\log(p)$ -value for different (τ_1, τ_2) , and $n = 1024$.

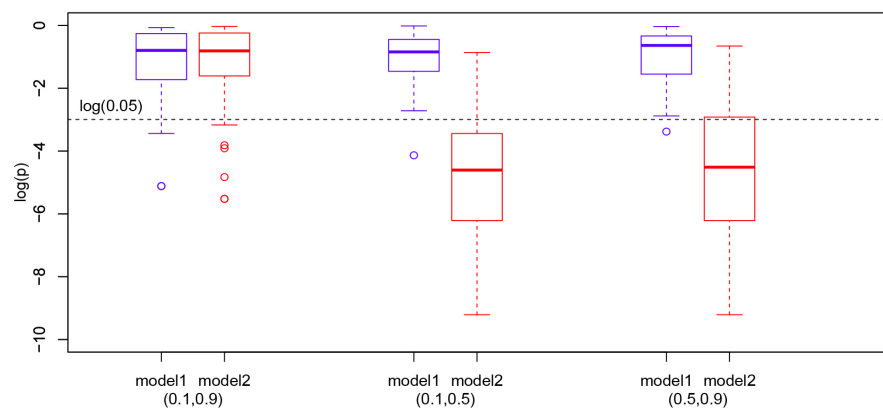


Figure 2. Boxplots of the estimated $\log(p)$ -value for different (τ_1, τ_2) , and $n = 2048$.

$P(x_t \leq q_{0.9}, x_{t+k} \leq q_{0.1})$, means that AR process with Gaussian innovations or t_1 distributed innovations is time-reversible. Also, when $(\tau_1, \tau_2) = (0.1, 0.5)$ or $(0.5, 0.9)$, these observations also reflect the fact that AR process with Gaussian innovations are time-reversible. However, for t_1 distributed innovations, this phenomenon only takes place for the extreme quantiles ($\tau_1 = 0.1, \tau_2 = 0.9$), does not hold for $\tau_1 = 0.5$ and $\tau_2 = 0.1$ or $\tau_2 = 0.9$, one of the important reason is that there exist a marked discrepancy between tail and central dependence structures when the innovations $\{\varepsilon_t\}$ of AR(1) process are non-Gaussian. From **Figure 2**, it can also be evidenced that AR process with Gaussian innovations is time reversible, while the case of t_1 distributed innovations is time reversible only for extreme quantiles. Above results indicate that the imaginary parts of CSDK are not zero suggesting time-irreversibility.

5. Conclusions

In this paper, we proposed a new statistic (3) for testing the specification of the cross-covariance matrix. The test statistic is constructed based on Frobenius norm of the difference between the sample cross-covariance matrix and the given matrix. The asymptotic properties of test statistic were obtained with the help of bootstrap scheme. By computing the empirical size and power of the pro-

posed test, the rationality of the test statistic was obtained. The advantage of the proposed statistic is twofold. First, through comparative study, we found that our empirical powers are clearly superior to others in detecting any deviation of the cross-covariance from the pre-assigned matrix. Second, there is no need to make complex derivations of the distribution of statistic T and only a few simulation studies we can obtain the performance of the test.

However, one challenge is to determine whether the test performs very well in the case where the data is high-dimensional, this will be our future work.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China (grant number: 11671416).

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Liu, J., Han, J.W., Zhang, Z.J. and Li, J. (2018) Target Detection Exploiting Covariance Matrix Structures in MIMO Radar. *Signal Processing*, **154**, 174-181. <https://doi.org/10.1016/j.sigpro.2018.07.013>
- [2] Yu, H. and Yang, J. (2001) A Direct LDA Algorithm for High-Dimensional Data with Application to Face Recognition. *Pattern Recognition*, **34**, 2067-2070. [https://doi.org/10.1016/S0031-3203\(00\)00162-X](https://doi.org/10.1016/S0031-3203(00)00162-X)
- [3] Mauchly, J.W. (1940) Significance Test for Sphericity of a Normal n-Variate Distribution. *The Annals of Mathematical Statistics*, **11**, 204-209. <https://doi.org/10.1214/aoms/1177731915>
- [4] Gupta, A.K. and Xu, J. (2006) On Some Tests of the Covariance Matrix under General Conditions. *Annals of the Institute of Statistical Mathematics*, **58**, 101-114. <https://doi.org/10.1007/s10463-005-0010-z>
- [5] Jiang, D.D., Jiang, T.F. and Yang, F. (2012) Likelihood Ratio Tests for Covariance Matrices of High-Dimensional Normal Distributions. *Journal of Statistical Planning and Inference*, **142**, 2241-2256. <https://doi.org/10.1016/j.jspi.2012.02.057>
- [6] Bai, Z.D., Jing, D.D., Zheng, S.R., *et al.* (2009) Corrections to LRT on Large-Dimensional Covariance Matrix by RMT. *The Annals of Statistics*, **37**, 3822-3840. <https://doi.org/10.1214/09-AOS694>
- [7] Niu, Z.Z., Hu, J.H., Bai, Z.D. and Gao, W. (2019) On LR Simultaneous Test of High-Dimensional Mean Vector and Covariance Matrix under Non-Normality. *Statistics and Probability Letters*, **145**, 338-344. <https://doi.org/10.1016/j.spl.2018.10.008>
- [8] Nagao, H. (1973) On Some Test Criteria for Covariance Matrix. *The Annals of Statistics*, **1**, 700-709. <https://doi.org/10.1214/aos/1176342464>
- [9] Cai, T.T. and Jiang, T.F. (2011) Limiting Laws of Coherence of Random Matrices with Applications to Testing Covariance Structure and Construction of Compressed Sensing Matrices. *The Annals of Statistics*, **39**, 1496-1525. <https://doi.org/10.1214/11-AOS879>

- [10] Gupta, A.K. and Bodnar, T. (2014) An Exact Test about the Covariance Matrix. *Journal of Multivariate Analysis*, **125**, 176-189.
<https://doi.org/10.1016/j.jmva.2013.12.007>
- [11] Chen, Y.T., Chou, R.Y. and Kuan, C.M. (2000) Testing Time Reversibility without Moment Restrictions. *Journal of Econometrics*, **95**, 199-218.
[https://doi.org/10.1016/S0304-4076\(99\)00036-6](https://doi.org/10.1016/S0304-4076(99)00036-6)
- [12] Ramsey, J.B. and Rothman, P. (1996) Time Irreversibility and Business Cycle Asymmetry. *Journal of Money, Credit and Banking*, **28**, 1-21.
<https://doi.org/10.2307/2077963>
- [13] Rothman, P. (1999) Higher-Order Residual Analysis for Simple Bilinear and Threshold Autoregressive Models with the TR Test. *Nonlinear Time Series Analysis of Economic and Financial Data*, **1**, 357-367.
https://doi.org/10.1007/978-1-4615-5129-4_16
- [14] Hinich, M.J. and Rothman, P. (1998) Frequency-Domain Test of Time Reversibility. *Macroeconomic Dynamics*, **2**, 72-88.
- [15] Cox, D.R. (1981) Statistical Analysis of Time Series: Some Recent Developments. *Scandinavian Journal of Statistics*, **8**, 110-111.
- [16] Dette, H., Hallin, M. and Kley, T. (2015) Of Copulas, Quantiles, Ranks and Spectra: An Approach to Spectral Analysis. *Bernoulli*, **21**, 781-831.
<https://doi.org/10.3150/13-BEJ587>
- [17] Bickel, P.J. and Freedman, D.A. (1981) Some Asymptotic Theory for the Bootstrap. *The Annals of Statistics*, **9**, 1196-1217. <https://doi.org/10.1214/aos/1176345637>
- [18] Sugiura, N. and Nagao, H. (1968) Unbiasedness of Some Test Criteria for Equality of One or Two Covariance Matrices. *The Annals of Mathematical Statistics*, **39**, 1686-1692.
<https://doi.org/10.1214/aoms/1177698150>
- [19] Anderson, T.W. (2003) An Introduction to Multivariate Statistical Analysis. 3rd Edition, John Wiley and Sons Press, New York.
- [20] Zhang, S.B. (2019) Bayesian Copula Spectral Analysis for Stationary Time Series. *Computational Statistics and Data Analysis*, **133**, 166-179.
<https://doi.org/10.1016/j.csda.2018.10.001>
- [21] Koenker, R. (2005) Quantile Regression. Cambridge University Press, Cambridge.
<https://doi.org/10.1017/CBO9780511754098>
- [22] Kley, T. (2016) Quantile-Based Spectral Analysis in an Object-Oriented Framework and a Reference Implementation in R: The Quantspec Package. *Journal of Statistical Software*, **70**, 1-27. <https://doi.org/10.18637/jss.v070.i03>