Journal of Multivariate Analysis 126 (2014) 25–52

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Limiting spectral distribution of renormalized separable sample covariance matrices when $p/n \to 0$

Lili Wang [a,b,*], Debashis Paul [b]

[a] Department of Mathematics, Zhejiang University, Hangzhou, 310027, China
[b] Department of Statistics, University of California, Davis, CA 95616, United States

ARTICLE INFO

Article history:
Received 2 August 2013
Available online 16 January 2014

AMS subject classifications:
60B20
62E20
60F05
60F15
62H99

Keywords:
Separable covariance
Limiting spectral distribution
Stieltjes transform
McDiarmid's inequality
Lindeberg principle
Wielandt's inequality

ABSTRACT

We are concerned with the behavior of the eigenvalues of renormalized sample covariance matrices of the form

$$C_n = \sqrt{\frac{n}{p}}\left(\frac{1}{n}A_p^{1/2}X_nB_nX_n^*A_p^{1/2} - \frac{1}{n}\mathrm{tr}(B_n)A_p\right)$$

as $p, n \to \infty$ and $p/n \to 0$, where X_n is a $p \times n$ matrix with i.i.d. real or complex valued entries X_{ij} satisfying $E(X_{ij}) = 0$, $E|X_{ij}|^2 = 1$ and having finite fourth moment. $A_p^{1/2}$ is a square-root of the nonnegative definite Hermitian matrix A_p, and B_n is an $n \times n$ nonnegative definite Hermitian matrix. We show that the empirical spectral distribution (ESD) of C_n converges a.s. to a nonrandom limiting distribution under the assumption that the ESD of A_p converges to a distribution F^A that is not degenerate at zero, and that the first and second spectral moments of B_n converge. The probability density function of the LSD of C_n is derived and it is shown that it depends on the LSD of A_p and the limiting value of $n^{-1}\mathrm{tr}(B_n^2)$. We propose a computational algorithm for evaluating this limiting density when the LSD of A_p is a mixture of point masses. In addition, when the entries of X_n are sub-Gaussian, we derive the limiting empirical distribution of $\{\sqrt{n/p}(\lambda_j(S_n) - n^{-1}\mathrm{tr}(B_n)\lambda_j(A_p))\}_{j=1}^p$ where $S_n := n^{-1}A_p^{1/2}X_nB_nX_n^*A_p^{1/2}$ is the sample covariance matrix and λ_j denotes the jth largest eigenvalue, when F^A is a finite mixture of point masses. These results are utilized to propose a test for the covariance structure of the data where the null hypothesis is that the joint covariance matrix is of the form $A_p \otimes B_n$ for $\otimes$ denoting the Kronecker product, as well as A_p and the first two spectral moments of B_n are specified. The performance of this test is illustrated through a simulation study.

© 2014 Elsevier Inc. All rights reserved.

1. Introduction

In this paper, we obtain the limiting spectral distribution (LSD) and a system of equations describing the corresponding Stieltjes transforms of renormalized sample covariance matrices of the form

$$C_n = \sqrt{\frac{n}{p}}\left(\frac{1}{n}A_p^{1/2}X_nB_nX_n^*A_p^{1/2} - \frac{1}{n}\mathrm{tr}(B_n)A_p\right) \tag{1.1}$$

* Corresponding author at: Department of Mathematics, Zhejiang University, Hangzhou, 310027, China.
E-mail addresses: sta.lesleywang@gmail.com, wanglili0307@163.com (L. Wang), debpaul@ucdavis.edu (D. Paul).

0047-259X/$ – see front matter © 2014 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jmva.2013.12.015

when $p, n \to \infty$ and $p/n \to 0$, where X_n has i.i.d. real or complex entries with zero mean, unit variance and uniformly bounded fourth moment. Throughout this paper, for any matrix M, we use M^* to denote the complex conjugate transpose of M if M is complex-valued and transpose of M if M is real-valued. When $p/n \to c \in (0, +\infty)$ as $n \to \infty$, the spectral properties of the separable sample covariance matrices, $S_n := n^{-1}A_p^{1/2}X_nB_nX_n^*A_p^{1/2}$ have been widely investigated under different assumptions on entries (e.g., Zhang [26], Paul and Silverstein [20], EL Karoui [9]). The name "separable" refers to the fact that the covariance matrix of the vectorized data matrix $Y_n = A_p^{1/2}X_nB_n^{1/2}$ has the separable covariance $A_p \otimes B_n$, where $\otimes$ denotes the Kronecker product between matrices. Under those circumstances, it is known that the spectral norm of $S_n - \mathbb{E}S_n$ does not converge to zero. However, if $p/n \to 0$, $\|S_n - \mathbb{E}S_n\| \xrightarrow{a.s.} 0$. When $A_p = I_p$, $B_n = I_n$ and $p, n \to \infty$ such that $p/n \to 0$, the behavior of empirical spectral distribution (ESD) of $\sqrt{n/p}\,(S_n - \mathbb{E}S_n) = \sqrt{n/p}(n^{-1}X_nX_n^* - I_p)$ is similar to that of a $p \times p$ Wigner matrix W_p, which has been verified by Bai and Yin [3]. Moreover, when $S_n = n^{-1}A_p^{1/2}X_nX_n^*A_p^{1/2}$, for i.i.d. real entries and under a finite fourth moment condition, Pan and Gao [19] and Bao [5] derived the LSD of $\sqrt{n/p}(n^{-1}A_p^{1/2}X_nX_n^*A_p^{1/2} - A_p)$, which coincides with that of a generalized Wigner matrix $p^{-1/2}A_p^{1/2}W_pA_p^{1/2}$ studied by Bai and Zhang [4]. Our work here extends the former result to a more general setting, namely, when B_n is an arbitrary $n \times n$ positive semi-definite matrix whose first two spectral moments converge to finite positive values as $n \to \infty$, and the entries of X_n are either real or complex. The strategy of the proof of this result is divided into three parts. We first assume that the entries of X_n are i.i.d. Gaussian and use a construction analogous to that in Pan and Gao [19] to obtain the form of the approximate deterministic equations describing the expected Stieltjes transforms, then use a result on the concentration of smooth functions of independent random elements to show that the Stieltjes transform concentrates around its mean in the general setting (without the restriction of Gaussianity), and finally utilize the Lindeberg principle to show that the expected Stieltjes transforms in the Gaussian and in the general case are asymptotically the same. In the process, we also prove the existence and uniqueness of the system of equations describing the Stieltjes transform for an arbitrary F^A, non-degenerate at zero. Further, we state a result characterizing the LSD, including the existence and shape of its density function, by following the approach in Bai and Zhang [4]. We also study the question of fluctuation of the eigenvalues of the sample covariance matrix $S_n := n^{-1}A_p^{1/2}X_nB_nX_n^*A_p^{1/2}$ itself when the ESD of A_p, say F^{A_p} and its limit F^A are finite mixtures of point masses. Specifically, we show that the empirical distribution of $\{\sqrt{n/p}(\lambda_j(S_n) - n^{-1}\mathrm{tr}(B_n)\lambda_j(A_p))\}_{j=1}^p$, where λ_j denotes the jth largest eigenvalue, converges a.s. to a mixture of rescaled semi-circle laws with mixture weights being the same as the weights corresponding to the point masses of F^A and the scaling factor depending on the limiting value of $n^{-1}\mathrm{tr}(B_n^2)$ and the atoms of F^A.

It should be noted that the data model of the form $Y_n = A_p^{1/2}X_nB_n^{1/2}$, where X_n has i.i.d. entries with zero mean and unit variance, relates very closely to the *separable covariance model* widely used in spatio-temporal data modeling, especially for modeling environmental data (e.g., Kyriakidis and Journel [11], Mitchell and Gumpertz [13], Fuentes [10], Li et al. [16]). The separable covariance model refers to the fact that for any p sampling locations in space, and any n observation times, the covariance of the corresponding data matrix can be expressed in the form $\Sigma_{n,p} = A_p \otimes B_n$. In that context, the rows of Y_n correspond to spatial locations while the columns represent the observation times. If furthermore, the process is Gaussian, which is often assumed in the literature, then the data matrix Y_n is exactly of the form $A_p^{1/2}X_nB_n^{1/2}$ where X_{ij}'s are i.i.d. $N(0, 1)$. Assuming a separable covariance structure, that the process is stationary in space, the sampling locations cover the entire spatial region under consideration fairly evenly, and the temporal variation has only short term dependence (not necessarily stationary), the covariance of the observed data can be expressed in the form $A_p \otimes B_n$ where A_p and B_n satisfy conditions 3–5 of our main result in this paper (Theorem 2.1). Moreover, if the sampling locations are on a spatial grid, then the matrix of eigenvectors of A_p is approximately the Fourier rotation matrix on $\mathbb{R}^p$ and the eigenvalues are approximately the Fourier transforms of the spatial autocovariance kernel evaluated at certain discrete frequencies related to the grid spacings.

There is a body of literature on the statistical inference for a separable covariance model, in particular about the tests for separability of the joint covariance of the data. Notable examples include Dutilleul [7], Lu and Zimmerman [17], Mitchell et al. [14,15], Fuentes [10], Roy and Khatree [21], Simpson [24] and Li et al. [16]. These tests typically assume joint Gaussianity of the data and often the derivation of the test statistic requires additional structural assumptions, e.g., stationarity of the spatial and temporal processes (Fuentes [10]). In addition, the estimation techniques often involve matrix inversions (Dutilleul [7], Mitchell et al. [15]) which become challenging if the dimensionality (either p or n) is large. We study the problem of tests involving the separable covariance structure under the framework $p, n \to \infty$ and $p/n \to 0$. Under this setting, $\|n^{-1}Y_nY_n^* - n^{-1}\mathrm{tr}(B_n)A_p\| \xrightarrow{a.s.} 0$ and hence we can infer about the spectral properties of A_p from that of the sample covariance matrix $n^{-1}Y_nY_n^*$. In particular, we propose to use the results derived here to construct test statistic for testing whether the space–time data follows a specific separable covariance model, where the null hypothesis is in terms of specification of A_p and the first two spectral moments of B_n. Let A_0, $\mathrm{tr}(B_0)$ and $\mathrm{tr}(B_0^2)$ be the specified values of A_p, $\mathrm{tr}(B_n)$ and $\mathrm{tr}(B_n^2)$ under the null hypothesis. Then this statistic measures the difference of the ESD of the matrix $\sqrt{n/p}(n^{-1}Y_nY_n^* - n^{-1}\mathrm{tr}(B_0)A_0)$, from the LSD of C_n described in (1.1), where the matrix X_n is assumed to have i.i.d. entries with zero mean and unit variance, $A_p = A_0$, $\mathrm{tr}(B_n) = \mathrm{tr}(B_0)$ and $\mathrm{tr}(B_n^2) = \mathrm{tr}(B_0^2)$. We also propose a Monte-Carlo method for determination of the cut-off values of the test for any given level of significance and analyze the behavior of the test through simulation. We also carry out a simulation study with different combinations of (p, n) to empirically assess the rate of convergence of the ESD under to the LSD as measured by the L^2 distance between these distributions.

Download English Version:

https://daneshyari.com/en/article/7546909

Download Persian Version:

https://daneshyari.com/article/7546909

Daneshyari.com