Research Article

Limiting Spectral Distribution of Large-Dimensional Sample Covariance Matrices Generated by the Periodic Autoregressive Model

Jin Zou and Dong Han

School of Mathematics Science, Shanghai Jiao Tong University, 800 Dongchuan RD. Minhang District, Shanghai, China

Correspondence should be addressed to Jin Zou; mathzj@alumni.sjtu.edu.cn

The explicit representation for the limiting spectral moments of sample covariance matrices generated by the periodic autoregressive model (PAR) is established. We propose to use the moment-constrained maximum entropy method to estimate the spectral density function. The experiments show that the maximum entropy spectral density function curve obtained based on the fourth-order limiting spectral moment can match histograms of the eigenvalues of the covariance matrices very well.

1. Introduction

Periodic time-series models are used for modeling stationary time series with periodic characteristics. Periodicity refers to wave or oscillating movements around long-term trends presented in the time series, such as the water sector, where the supply of water varies with seasonal rainfall. In the economic time series, it is often accompanied by some cycle or quasi-periodic fluctuations, such as day and night changes of urban electricity consumption, seasonal changes of railway passenger transport, and repeated acceleration, interruption, stagnation, and recovery of economic growth; see Franses and Paap [1], Bell et al. [2], and Aliat and Hamdi [3]. The periodic autoregressive (PAR) model is a generalisation of the classical autoregressive (AR) model by allowing the parameters to vary with the variables over a period:

$$\mathbf{X}_{ip+v} = \sum_{k=1}^{q(v)} \alpha_k(v)\mathbf{X}_{ip+v-k} + \boldsymbol{\varepsilon}_{ip+v}. \tag{1}$$

For a given v and a predetermined value p, the random variable X_{ip+v} denotes the value in the vth quarter of the ith cycle, where $v \in \{1,\ldots,p\}$ and $i \in \{1,\ldots,n\}$. In the ith cycle, X_{ip+v} is an AR process. The model order is given by $q(v)$, and

$\alpha_k(v), k = 1,\ldots,q(v)$, represents the AR model coefficients for the vth quarter. The error process $\varepsilon = \{\varepsilon_{ip+v}\}$ corresponds to a periodic white noise, with $E(\varepsilon_{ip+v}) = 0$ and $\operatorname{var}(\varepsilon_{ip+v}) = \sigma^2(v) > 0$. Let $\mathbf{X}_i = (\mathbf{X}_{ip+1},\ldots,\mathbf{X}_{ip+p})^\top$, where the transpose (indicated by $\top$) of a row vector is a column vector, and we define the large-dimensional sample covariance matrices as $\mathbf{B}_n = (1/n)\sum_{i=1}^n \mathbf{X}_i\mathbf{X}_i^\top$, with the periodic length p tends to infinity in proportion to the sample size n, namely, $p/n \longrightarrow c \in (0,\infty)$. Sample covariance matrices are very important in multivariate statistical inference since many test statistics are defined by their eigenvalues or functions.

Random matrix theory (RMT) originated from the study of the energy levels of a large number of particles in quantum mechanics. Many of the laws of mathematical physics were discovered through numerical studies. In the late 1950s, E. P. Wigner formulated the problem in terms of the empirical distribution of random matrices [4, 5], which led to the study of the semicircular law of Gaussian matrices. Since then, RMT has formed an active branch in modern probability theory. The LSD of the sample covariance matrix of a large-dimensional random matrix composed of independent random variables, the MP law, was first established by Marcenko and Pastur [6]. Further research efforts were

conducted to estimate the LSD of a product of two random matrices. Yin and Krishnaiah [7] proved the existence of LSD for the matrix sequence $\mathbf{S}_n\mathbf{T}_n$ when the two matrix sequences $\mathbf{S}_n$ and $\mathbf{T}_n$ are standard Wishart matrices and positive definite matrices, respectively. In particular, Bai et al. [8] obtained a display representation of the LSD of the matrix sequence $\mathbf{S}_n\mathbf{T}_n$ when the two matrix sequences $\mathbf{S}_n$ and $\mathbf{T}_n$ are the sample covariance matrix and the Wigner matrix, respectively. To relax the independence of the entries of $\mathbf{X}_i$, Silverstein [9] considered the case of $\mathbf{X}_i = \mathbf{M}\mathbf{x}_i$, where $\mathbf{M}$ is a nonnegative definite matrix and $\mathbf{x}_i$ consists of independent and identically distribution (i.i.d.) entries, and the sample covariance matrices take the form of $\mathbf{B}_n = (1/n)\sum_{i=1}^{n} \mathbf{M}\mathbf{x}_i\mathbf{x}_i^{\top}\mathbf{M}^{\top}$. Random matrices of the form $\mathbf{B}_n$ were extensively investigated by many researchers, including Yin [10], Silverstein [9], Bai and Zhou [11], Jin et al. [12], Zhang et al. [13] and Yao [14]. Silverstein's important work [9] aimed to relax the independence structure between $\mathbf{X}_i$ coordinates and considered random vectors of the form $\mathbf{M}\mathbf{x}_i$, assuming that the spectral norm of the matrix sequence $\mathbf{M}\mathbf{M}^{\top}$ is bounded and that the ESD of the matrix sequence $\mathbf{M}\mathbf{M}^{T}$ is convergent. Silverstein proved that the LSD of the sample covariance matrix $\mathbf{B}_n$ exists, given by the LSD of equation representation of the Stieltjes transformation.

This paper investigates the LSD of large-dimensional sample covariance matrices generated by the PAR model, and the sample covariance matrices take the form of $\mathbf{B}_n$. A similar ARMA-type processes problem has been considered by Bai and Zhou [11], Jin et al. [12], Yao [14], etc. It follows that, in the ARMA-type processes, the covariance matrices $\mathbf{M}\mathbf{M}^{\top}$ equals to the Toeplitz matrix, according to the fundamental eigenvalue distribution theorem for Toeplitz-type matrices [15], the ESD of $\mathbf{M}\mathbf{M}^{\top}$ weakly converges to the nonrandom distribution H. For some situations, the ESD of $\mathbf{M}\mathbf{M}^{\top}$ have explicit form; then, the LSD for the sample covariance matrix $\mathbf{B}_n$ can be solved explicitly. For the random matrix generated by the first-order vector autoregressive (VAR(1)) model and the first-order vector moving average (VMA(1)) model, Jin et al. [12] and Bai and Zhou [11] obtained a display representation of this LSD. However, in our first-order variable-coefficient PAR process, we consider the process be interrupted by certain factors that the odd terms and the even terms of the coefficient $\{\alpha_v\}$ converge to α and β, respectively, and that $|\alpha| \in (0,1)$ and $|\beta| \in (0,1)$. The mixing matrix $\mathbf{Q}$ is a lower triangular matrix whose main diagonal elements are all 1 and subdiagonal elements differ from each other. When $\alpha = \beta$, the limiting form of the matrices $\mathbf{Q}\mathbf{Q}^{\top}$ are Toeplitz matrices, and the eigenvalues of Toeplitz matrices have been proved by Gregory [16]; when $\alpha \neq \beta$, the ESD of $\mathbf{Q}\mathbf{Q}^{\top}$ has no explicit form, see [17, 18], and the structure of the limiting matrices $\mathbf{Q}^{-1}(\mathbf{Q}^{-1})^{\top}$ are very complex. Our main contributions in this paper include the following. (i) For the first-order variable-coefficient PAR model, the explicit forms of limiting spectral moments for $\mathbf{Q}^{-1}(\mathbf{Q}^{-1})^{\top}$ are given; through verifying the Carleman condition, we derive the explicit spectral moments of $(1/n)\sum_{i=1}^{n} \mathbf{X}_i\mathbf{X}_i^{\top}$. (ii) A framework of the maximum entropy-based method is provided for estimate the LSD.

The paper is organized as follows. In Section 2, we employ the moment method to deduce the spectral density of sample covariance matrices. In Section 3, we use the maximum entropy method to find the approximate solution and analyze the convergence of the maximum entropy solution. We simulate the PAR model to obtain the periodic sample covariance matrices, and the experiment shows that the maximum entropy spectral density function curve can match histograms of the eigenvalues of the covariance matrices very well.

2. Methodology and Main Results

In this section, we first derive the explicit spectral moments form of $\mathbf{Q}^{-1}(\mathbf{Q}^{-1})^{\top}$; subsequently, we verify the spectral moments statistics satisfying the Carleman condition, and then, we employ the moment method to deduce the spectral density of sample covariance.

We consider first-order PAR model:

$$\mathbf{X}_{ip+v} = \alpha_v\mathbf{X}_{ip+v-1} + \boldsymbol{\varepsilon}_{ip+v}. \tag{2}$$

Let

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ -\alpha_1 & 1 & 0 & \cdots & 0 \\ 0 & -\alpha_2 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & -\alpha_{p-1} & 1 \end{bmatrix},$$

$$\mathbf{X}_i = \left(\mathbf{X}_{ip+1}, \ldots, \mathbf{X}_{ip+p}\right)^{\top},$$

$$\boldsymbol{\varepsilon}_i = \left(\boldsymbol{\varepsilon}_{ip+1}, \ldots, \boldsymbol{\varepsilon}_{ip+p}\right)^{\top}. \tag{3}$$

Then, the first-order PAR model can be defined as

$$\mathbf{Q}\mathbf{X}_i = \boldsymbol{\varepsilon}_i \Rightarrow \mathbf{X}_i^{\top}\mathbf{Q}^{\top} = \boldsymbol{\varepsilon}_i^{\top}, \quad i = 1, \ldots, n, \tag{4}$$

where $\Rightarrow$ denotes the equation on the left can push the equation on the right. From equation (3), $\mathbf{Q}$ is a lower triangular matrix whose main diagonal elements are all 1, and therefore, the matrix $\mathbf{Q}$ is invertible. Hence,

$$\mathbf{Q}\mathbf{X}_i\mathbf{X}_i^{\top}\mathbf{Q}^{\top} = \boldsymbol{\varepsilon}_i\boldsymbol{\varepsilon}_i^{\top} \Rightarrow \mathbf{X}_i\mathbf{X}_i^{\top} = \mathbf{Q}^{-1}\boldsymbol{\varepsilon}_i(\boldsymbol{\varepsilon}_i)^{\top}(\mathbf{Q}^{-1})^{\top}$$

$$\Rightarrow \frac{1}{n}\sum_{i=1}^{n} \mathbf{X}_i\mathbf{X}_i^{T} = \mathbf{Q}^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} \boldsymbol{\varepsilon}_i(\boldsymbol{\varepsilon}_i)^{\top}\right)(\mathbf{Q}^{-1})^{\top}. \tag{5}$$

The sample covariance matrices $(1/n)\sum_{i=1}^{n} \mathbf{X}_i(\mathbf{X}_i)^{\top}$ and $\mathbf{Q}^{-1}((1/n)\sum_{i=1}^{n} \boldsymbol{\varepsilon}_i(\boldsymbol{\varepsilon}_i)^{\top})(\mathbf{Q}^{-1})^{\top}$ have the same ESD. Let $\mathbf{S} = (1/n)\sum_{i=1}^{n} \boldsymbol{\varepsilon}_i(\boldsymbol{\varepsilon}_i)^{\top}$, and the ESD of $\mathbf{Q}^{-1}((1/n)\sum_{i=1}^{n} \boldsymbol{\varepsilon}_i (\boldsymbol{\varepsilon}_i)^{\top})(\mathbf{Q}^{-1})^{\top}$ is the same as that of $\mathbf{Q}^{-1}(\mathbf{Q}^{-1})^{\top}\mathbf{S}$. Next, we establish the existence of the LSD of the covariance matrix sequence $(1/n)\sum_{i=1}^{n} \mathbf{X}_i(\mathbf{X}_i)^{\top}$.

Theorem 1. *Let $\mathbf{X}_{ip+v}$ be generated by a first-order PAR model in equation (2), and the odd terms and the even terms of the coefficient $\{\alpha_v\}$ converge to α and β, respectively, and $|\alpha| \in (0,1), |\beta| \in (0,1)$. The error $\boldsymbol{\varepsilon}_{ip+v}$ is independent*

random variables satisfying for any δ, $(1/\delta^2 np)\sum_{i,t} E(|\varepsilon_{i,t}|^2 I)(|\varepsilon_{i,t}| \geq \delta\sqrt{n}) \longrightarrow 0$; furthermore, we assume that $p/n \longrightarrow c \in (0, \infty)$; then, the spectral moments of sample covariance matrix $(1/n)\sum_{i=1}^{n} \mathbf{X}_i(\mathbf{X}_i)^\top$ tend to

$$\beta\left(\frac{1}{n}\sum_{i=1}^{n}\mathbf{X}_i(\mathbf{X}_i)^\top\right)_k = \int \lambda^k dF(\lambda) \xrightarrow{a.s.} \sum_{s=1}^{k} c^{k-s}$$
$$\cdot \sum_{\substack{i_1+\cdots+i_s=k-s+1 \\ i_1+\cdots+si_s=k}} \frac{k!}{s!} \prod_{m=1}^{s} \frac{\mu_m^{i_m}}{i_m!}, \qquad (6)$$

where μ_m is the limit of spectral moments of $\mathbf{Q}^{-1}(\mathbf{Q}^{-1})^\top$, which have an explicit form:

$$\mu_m = \frac{(\alpha+\beta)^2}{4\alpha\beta(1-\alpha\beta)^{2m-1}} \sum_{j=0}^{m-1} \left(C_{m-1}^{j}\right)^2 (\alpha\beta)^j$$
$$- \frac{(\alpha-\beta)^2}{4\alpha\beta(1+\alpha\beta)^{2m-1}} \sum_{j=0}^{m-1} \left(C_{m-1}^{j}\right)^2 (-\alpha\beta)^j. \qquad (7)$$

Proof. The proof is divided into two steps. We first show the existence of the limit of spectral moments which is defined as $\lim_{p\to\infty}(1/p)\mathrm{tr}((\mathbf{Q}^{-1})^m((\mathbf{Q}^{-1})^m)^\top)$; then, we verify the Carleman condition. From equation (3), the inverse of matrix $\mathbf{Q}$ can be written as

$$\mathbf{Q}^{-1} = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ \alpha_1 & 1 & 0 & \cdots & 0 \\ \alpha_1\alpha_2 & \alpha_2 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \prod_{k=1}^{p-1}\alpha_k & \prod_{k=2}^{p-1}\alpha_k & \cdots & \prod_{k=p-1}^{p-1}\alpha_k & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ a_{21} & 1 & 0 & \cdots & 0 \\ a_{31} & a_{32} & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ a_{p1} & \cdots & \cdots & a_{p,p-1} & 1 \end{bmatrix}, \qquad (8)$$

where a_{ij} can be written as

$$a_{ij} = \prod_{k=j}^{i-1} \alpha_k, \quad j < i, a_{ii} = 1, i, j = 1, 2, \ldots, p-1. \qquad (9)$$

Let

$$c_m^{(i)} = C_{m-1+i-1}^{i-1} = \frac{(m-1+i-1)!}{(i-1)!(m-1)!}. \qquad (10)$$

Then,

$$\left(\mathbf{Q}^{-1}\right)^m = \begin{bmatrix} 1 & 0 & \cdots & \cdots & 0 \\ c_m^{(2)}a_{21} & 1 & 0 & \cdots & 0 \\ c_m^{(3)}a_{31} & c_m^{(2)}a_{32} & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ c_m^{(p)}a_{p1} & \cdots & \cdots & c_m^{(2)}a_{p,p-1} & 1 \end{bmatrix}. \qquad (11)$$

Furthermore, we obtain that

$$\mathrm{tr}\left(\mathbf{Q}^{-1}\left(\mathbf{Q}^{-1}\right)^\top\right)^m = \mathrm{tr}\left(\left(\mathbf{Q}^{-1}\right)^m\left(\left(\mathbf{Q}^{-1}\right)^\top\right)^m\right)$$
$$= \mathrm{tr}\left(\left(\mathbf{Q}^{-1}\right)^m\left(\left(\mathbf{Q}^{-1}\right)^m\right)^\top\right). \qquad (12)$$

Applying the trace operator, we have

$$\frac{1}{p}\mathrm{tr}\left(\mathbf{Q}^{-1}\left(\mathbf{Q}^{-1}\right)^\top\right)^m = 1 + \sum_{N=2}^{p}\left(\frac{1}{p}\sum_{k=N}^{p}\left(c_m^{(N)}\right)^2 a_{k\,k-(N-1)}^2\right). \qquad (13)$$

Based on the assumption that $\{\alpha_{2v+1}\} \longrightarrow \alpha$, $\{\alpha_{2v+1}\} \longrightarrow \beta$, as $v \longrightarrow \infty$, $|\alpha| \in (0,1), |\beta| \in (0,1)$, we obtain

$$\frac{1}{p}\sum_{k=N}^{p}\left(c_m^{(N)}\right)^2 a_{k\,k-(N-1)}^2 \longrightarrow \left(c_m^{(N)}\right)^2\left(\frac{(-1)^{N+1}+1}{2}\alpha^{N-1}\beta^{N-1}\right.$$
$$\left. + \frac{(-1)^N+1}{2}\frac{\alpha^N\beta^{N-2}+\alpha^{N-2}\beta^N}{2}\right), \quad \text{as } p \longrightarrow \infty, \qquad (14)$$

and therefore,

$$\frac{1}{p}\mathrm{tr}\left(\left(\mathbf{Q}^{-1}\right)^m\left(\left(\mathbf{Q}^{-1}\right)^m\right)^\top\right) \longrightarrow \frac{(\alpha+\beta)^2}{4\alpha\beta}\sum_{k=0}^{\infty}\left(c_m^{(k+1)}\right)^2(\alpha\beta)^k$$
$$- \frac{(\alpha-\beta)^2}{4\alpha\beta}\sum_{k=0}^{\infty}\left(c_m^{(k+1)}\right)^2(-\alpha\beta)^k. \qquad (15)$$

Let

$$h(x) = \sum_{k=0}^{\infty}\left(c_m^{(k+1)}\right)^2 x^k. \qquad (16)$$

Then,

$$\frac{1}{p}\mathrm{tr}\left(\left(\mathbf{Q}^{-1}\right)^m\left(\left(\mathbf{Q}^{-1}\right)^m\right)^\top\right) \longrightarrow \frac{(\alpha+\beta)^2}{4\alpha\beta}h(\alpha\beta) - \frac{(\alpha-\beta)^2}{4\alpha\beta}h(-\alpha\beta). \qquad (17)$$

The following section wishes to obtain a simplified form of the function $h(x)$. From equations (10) and (16), we have

$$h(x) = \sum_{k=0}^{\infty}\left(\frac{(m-1+k)!}{k!(m-1)!}\right)^2 x^k. \qquad (18)$$

Define

$$f(x) = \frac{1}{((m-1)!)^2}\sum_{k=0}^{\infty}\left((k+1)(k+2),\ldots,(m-1+k)x^{m-1-k}\right). \qquad (19)$$

Then, $h(x)$ can be rewritten as

$$h(x) = \frac{d^{m-1}f(x)}{dx^{m-1}}. \qquad (20)$$

Let

$$g(x) = \frac{1}{((m-1)!)^2} \sum_{k=0}^{\infty} x^{m-1+k} = \frac{1}{((m-1)!)^2} \frac{x^{m-1}}{1-x}. \quad (21)$$

Therefore, $f(x)$ can have a simplified representation as

$$f(x) = \frac{d^{m-1}g(x)}{dx^{m-1}} \cdot x^{m-1}. \quad (22)$$

By the derivation operations,

$$\frac{d^{m-1}g(x)}{dx^{m-1}} = \frac{1}{((m-1)!)^2} \sum_{i=0}^{m-1} C_{m-1}^i$$
$$\cdot \left(\prod_{k=1}^{i} (m-k) \right) x^{m-1-i} (m-1-i)! (1-x)^{-(m-i)}$$
$$= \frac{1}{(m-1)!} \sum_{i=0}^{m-1} C_{m-1}^i x^{m-1-i} (1-x)^{-(m-i)}$$
$$= \frac{1}{(m-1)!(1-x)^m}, \quad (23)$$

we obtain that

$$f(x) = \frac{x^{m-1}}{(m-1)!(1-x)^m},$$

$$\frac{d^{m-1}f(x)}{dx^{m-1}} = \frac{1}{(m-1)!} \sum_{i=0}^{m-1} C_{m-1}^i \left(\prod_{k=1}^{i} (m-k) \right) x^{m-1-i}$$
$$\cdot \left(\prod_{r=0}^{m-i-2} (m+r) \right) (1-x)^{-(2m-1-i)}. \quad (24)$$

By partially replacing x in equation (24) with y, we define a new function

$$\tilde{f}(x, y) = \frac{1}{(m-1)! x^{m-1} (1-y)^{2m-1}} \sum_{i=0}^{m-1} C_{m-1}^i$$
$$\cdot \left(\prod_{k=m-i}^{2m-i-2} k \right) x^{2m-2-i} (1-y)^i. \quad (25)$$

Define

$$H(x, y) = \sum_{i=0}^{m-1} C_{m-1}^i (m-i)$$
$$\cdot (m-i+1), \ldots, (2m-i-2) x^{2m-2-i} (1-y)^i,$$

$$M(x, y) = \sum_{i=0}^{m-1} C_{m-1}^i x^{2m-2-i} (1-y)^i = x^{m-1} (1-y+x)^{m-1}, \quad (26)$$

by which

$$H(x, y) = \frac{d^{m-1}M(x, y)}{dx^{m-1}} \cdot x^{m-1},$$

$$\frac{d^{m-1}M(x, y)}{dx^{m-1}} = \sum_{i=0}^{m-1} C_{m-1}^i \left(\prod_{k=1}^{i} (m-k) \right) x^{m-1-i}$$
$$\cdot \left(\prod_{k=1}^{m-1-i} (m-k) \right) (1-y+x)^i$$
$$= \sum_{i=0}^{m-1} \left(C_{m-1}^i \right)^2 x^{m-1-i} (m-1)! (1-y+x)^i, \quad (27)$$

so

$$H(x, y) = x^{m-1} \sum_{i=0}^{m-1} \left(C_{m-1}^i \right)^2 x^{m-1-i} (m-1)! (1-y+x)^i. \quad (28)$$

Then,

$$\tilde{f}(x, y) = \frac{x^{m-1}}{(m-1)! x^{m-1} (1-y)^{2m-1}} \sum_{i=0}^{m-1} \left(C_{m-1}^i \right)^2 x^{m-1-i}$$
$$\cdot (m-1)! (1-y+x)^i$$
$$= \frac{x^{m-1}}{(1-y)^{2m-1}} \sum_{i=0}^{m-1} \left(C_{m-1}^i \right)^2 x^{-i} (1-y+x)^i. \quad (29)$$

We use x to take place of y and obtain that

$$h(x) = \frac{x^{m-1}}{(1-x)^{2m-1}} \sum_{i=0}^{m-1} \left(C_{m-1}^i \right)^2 x^{-i}$$
$$= \frac{1}{(1-x)^{2m-1}} \sum_{j=0}^{m-1} \left(C_{m-1}^j \right)^2 x^j. \quad (30)$$

Therefore,

$$\mu_m = \lim_{p \longrightarrow \infty} \frac{1}{p} \text{tr}\left(\left(\mathbf{Q}^{-1} \right)^m \left(\left(\mathbf{Q}^{-1} \right)^m \right)^\top \right)$$
$$= \frac{(\alpha+\beta)^2}{4\alpha\beta(1-\alpha\beta)^{2m-1}} \sum_{j=0}^{m-1} \left(C_{m-1}^j \right)^2 (\alpha\beta)^j$$
$$- \frac{(\alpha-\beta)^2}{4\alpha\beta(1+\alpha\beta)^{2m-1}} \sum_{j=0}^{m-1} \left(C_{m-1}^j \right)^2 (-\alpha\beta)^j. \quad (31)$$

Next, we verify the Carleman condition $\sum_{k=1}^{\infty} \mu_{2k}^{-(1/2k)} = \infty$.
From equation (30), it follows that

$$\left(\mu_{2k}\right)^{-(1/2k)} = \left(\frac{(\alpha+\beta)^2}{4\alpha\beta(1-\alpha\beta)^{2m-1}}\sum_{j=0}^{m-1}\left(C_{m-1}^{j}\right)^2(\alpha\beta)^j - \frac{(\alpha-\beta)^2}{4\alpha\beta(1+\alpha\beta)^{2m-1}}\sum_{j=0}^{m-1}\left(C_{m-1}^{j}\right)^2(-\alpha\beta)^j\right)^{-(1/2k)}. \tag{32}$$

Then, applying the equality $C_{2n}^{n} = \sum_{j=0}^{n}\left(C_{n}^{j}\right)^2$ and $|\alpha| \in (0,1), |\beta| \in (0,1)$, we obtain

$$\left(\mu_{2k}\right)^{-(1/2k)} \geq \left(\frac{(|\alpha|+|\beta|)^2}{2|\alpha||\beta|(1-|\alpha\beta|)^{4k-1}}C_{4k-2}^{2k-1}\right)^{-(1/2k)}. \tag{33}$$

From Stirling's formula $n! = (1+o(1))\sqrt{2\pi n}n^n e^{-n}$, we get $n! \geq n^n e^{-n}$. Then, $C_{4k-2}^{2k-1} = (\prod_{j=0}^{2k-2}(4k-2-j)/(2k-1)!) \leq ((4k-2)^{2k-1}/(2k-1)!) \leq ((4k-2)e/(2k-1))^{2k-1}$; applying this inequality, $\left(\mu_{2k}\right)^{-(1/2k)}$ has a finite lower bound:

$$\left(\mu_{2k}\right)^{-(1/2k)} \geq \frac{(1-|\alpha\beta|)^2}{2e}. \tag{34}$$

Then,

$$\sum_{k=1}^{\infty}\mu_{2k}^{-(1/2k)} \geq \sum_{k=1}^{\infty}\left(\frac{(1-|\alpha\beta|)^2}{2e}\right) = \infty. \tag{35}$$

With the moment convergence theorem, the spectral distribution of $\mathbf{Q}^{-1}(\mathbf{Q}^{-1})^{\top}$ tends to a nonrandom probability distribution H. Let $\mathbf{S} = (1/n)\sum_{i=1}^{n}\varepsilon_i(\varepsilon_i)^{\top}$ and $p/n \longrightarrow c \in (0,\infty)$; by the important Theorem 2.10 of Bai [19], the ESD of $(1/n)\sum_{i=1}^{n}\mathbf{X}\mathbf{X}^{\top}$ tends to a nonrandom limit distribution in probability (or almost surely). The spectral moments of $(1/n)\sum_{i=1}^{n}\mathbf{X}\mathbf{X}^{\top}$ satisfy the following equation of Yin [10]:

$$\beta\left(\frac{1}{n}\sum_{i=1}^{n}XX^{\top}\right)_k = \int\lambda^k dF(\lambda) \xrightarrow{\text{a.s.}} \sum_{s=1}^{k}c^{k-s}\sum_{\substack{i_1+\cdots+i_s=k-s+1\\i_1+\cdots+si_s=k}}\frac{k!}{s!}\prod_{m=1}^{s}\frac{u_m^{i_m}}{i_m!}. \tag{36}$$

From the moment convergence theorem, we can obtain that the ESD of $(1/n)\sum_{k=1}^{n}\mathbf{X}\mathbf{X}^{\top}$ converges with probability close to 1 to a nonrandom distribution, determined by the spectral moment statistic. □

Remark 1. By the proof of Theorem 1, the limiting spectral moments of $\mathbf{Q}^{-1}(\mathbf{Q}^{-1})^{\top}$ depend only on the limiting coefficient α and β. When $\alpha = \beta$, the limiting form of the matrices $\mathbf{Q}\mathbf{Q}^{\top}$ are Toeplitz matrices, and the eigenvalues of Toeplitz matrices have been proved, see [16]. When $\alpha \neq \beta$, the limiting matrices of $\mathbf{Q}\mathbf{Q}^{\top}$ are tridiagonal matrices and take the form like the following:

$$\begin{bmatrix} -\alpha & 1+\alpha^2 & -\beta & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & -\beta & 1+\beta^2 & -\alpha & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & -\alpha & 1+\alpha^2 & -\beta & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & -\beta & 1+\beta^2 & -\alpha & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}. \tag{37}$$

The eigenvalues of this limiting matrix have no explicit form, see [18] and [17]. The limiting form of the matrices $\mathbf{Q}^{-1}(\mathbf{Q}^{-1})^{\top}$ are very complex, and we obtain the explicit spectral moments of $\mathbf{Q}^{-1}(\mathbf{Q}^{-1})^{\top}$.

3. Estimating Spectral Density Function with Maximum Entropy

Unfortunately, there are no explicit LSD forms of this sample covariance matrices $\mathbf{B}_n$. In this section, with the spectral moments of $\beta((1/n)\sum_{i=1}^{n}\mathbf{X}\mathbf{X}^{\top})_k$, we employ the maximum entropy method to estimate the LSD. Mead and Papanicolaou [20] studied the maximum entropy method to solve the uncertainty problem. Granziol et al. [21] demonstrated the equivalence between the MaxEnt principle and the optimization of variables with constraints. Gao and Han [22] used the maximum entropy method to solve nonlinear expectation. Every possible event has a probability of occurring, information entropy is related to all possibilities, and a system with the highest information entropy is most likely, when all known information, taken into account. In mathematics, the variable is generally estimated by maximising the Shannon entropy of the information of the moments of the random variable of the density function:

$$S = -\int_{\Omega}p(x)\ln p(x)dx, \tag{38}$$

where Ω is the support of the distribution. In this paper, the sample covariance matrices $\mathbf{Q}^{-1}((1/n)\sum_{i=1}^{n}\varepsilon_i(\varepsilon_i)^{\top})(\mathbf{Q}^{-1})^{\top}$ is nonnegative definite, so the estimated limiting distribution should be supported on $\Omega = \mathbb{R}^+ \cup \{0\}$.

Given the first $N+1$ moments $\beta_0, \beta_1, \ldots, \beta_N$, the idea of determining the density function of a probability distribution based on the principle of maximum entropy is to find the density function $p(x)$ such that

$$\int_{\Omega}x^k p(x)dx = \beta_k, \quad k = 0, 1, \ldots, N, \tag{39}$$

where $\beta_0 = 1$.

Lagrange multiplier is a method for finding the extremes of a multivariate function under a set of constraints. With the help of the Lagrange multiplier method, the entropy function is defined as

$$H \equiv \int_{\Omega}\left[-p(x)\ln p(x) + \sum_{k=0}^{N}\lambda_k x^k p(x)\right]dx - \sum_{k=0}^{N}\lambda_k\beta_k. \tag{40}$$

When the first-order derivative of the function H with respect to λ_k and $p(x)$ are equal to zero, the constraint defined in equation (39) is returned; then, the first-order derivative of the function do not provide any additional information. The second of these derivatives, however, evaluate to

$$\ln p_N(x) = -1 + \sum_{k=0}^{N} \lambda_k x^k. \tag{41}$$

The general solution of equation (41) is

$$p_N(x) = e^{-1+\sum_{k=0}^{N} \lambda_k x^k} = e^{\sum_{k=0}^{N} \lambda_k x^k}, \tag{42}$$

where we have set $(\lambda_0 - 1) \longrightarrow \lambda_0$.

We obtain the maximum entropy solution by solving for the Lagrange multiplier λ_k. To do this, we must solve the following nonlinear equation:

$$\begin{aligned} \int_\Omega e^{\lambda_0 + \lambda_1 x + \cdots + \lambda_k x^N} dx &= \beta_0 \\ \int_\Omega x e^{\lambda_0 + \lambda_1 x + \cdots + \lambda_k x^N} dx &= \beta_1 \\ &\vdots \\ \int_\Omega x^k e^{\lambda_0 + \lambda_1 x + \cdots + \lambda_k x^N} dx &= \beta_k. \end{aligned} \tag{43}$$

When $N \leq 2$, we can obtain a display representation of the Lagrange multiplier. Taking $N = 2$ as an example, from equation (36), we have

$$\mu_1 = \frac{2 + \alpha^2 + \beta^2}{2(1 - \alpha^2 \beta^2)},$$

$$\mu_2 = \frac{(\alpha + \beta)^2 (1 + \alpha\beta)^4 - (\alpha - \beta)^2 (1 - \alpha\beta)^4}{4\alpha\beta(1 - \alpha^2 \beta^2)^3}, \tag{44}$$

$$\beta_1 = \mu_1,$$

$$\beta_2 = y\mu_1^2 + \mu_2.$$

Based on equation (40), we obtain $\lambda_0, \lambda_1, \lambda_2$ as

$$\lambda_0 = -\left(\beta_1^2 + \ln\sqrt{2\pi(\beta_2 - \beta_1^2)} + 1\right),$$

$$\lambda_1 = \frac{\beta_1}{\beta_2 - \beta_1^2}, \tag{45}$$

$$\lambda_2 = -\frac{1}{2(\beta_2 - \beta_1^2)}.$$

Then,

$$p(x) = \exp(\lambda_0 + \lambda_1 x + \lambda_2 x^2). \tag{46}$$

When $N \geq 3$, we cannot obtain a display solution for the Lagrange multiplier λ_k, but at this point, the numerical solution for λ_k can be calculated using numerical optimization algorithms. From equation (40), the analytic expressions of the gradient is

$$\frac{\partial H}{\partial \lambda_j} = \beta_j - \int_\Omega x^j \exp\left(\sum_{i=0}^{m} \lambda_i x^i\right) dx. \tag{47}$$

If we denote the maximum entropy solution obtained by the method based on moments as

$$\tilde{\mu}_k \equiv \int_\Omega x^k e^{\lambda_0 + \lambda_1 x + \cdots + \lambda_k x^k} dx, \tag{48}$$

then the Hessian matrix is given by

$$G = \begin{bmatrix} \tilde{\mu}_0 & \tilde{\mu}_1 & \cdots & \tilde{\mu}_k \\ \tilde{\mu}_1 & \tilde{\mu}_2 & \cdots & \tilde{\mu}_{k+1} \\ \tilde{\mu}_2 & \tilde{\mu}_3 & \cdots & \tilde{\mu}_{k+2} \\ \vdots & \vdots & \vdots & \vdots \\ \tilde{\mu}_k & \tilde{\mu}_{k+1} & \cdots & \tilde{\mu}_{k+k} \end{bmatrix}. \tag{49}$$

$\tilde{\mu}_k$ is the kth moment of the simulated distribution, and the Newton iteration is done by searching λ_k until the termination criterion $|\tilde{\mu}_k - \mu_k| < \varepsilon$ holds. We obtain the spectral density $p(x)$ through Algorithm 1.

We can obtain that the maximum entropy solution $p_N(x)$ is weakly convergent, i.e.,

$$\lim_{N \longrightarrow +\infty} \int_\Omega p_N(x)\ln p_N(x)dx = \int_\Omega p(x)\ln p(x)dx. \tag{50}$$

In our simulations, we use the Newton method from the Python-based SciPy package to solve for minimisation $|\tilde{\mu}_k - \mu_k|$. It computes the gradient within a ending accuracy ε as well as the Hessian matrix. To give the Hessian matrix a faster rate of convergence, we symmetrize it and add a disturbance of intensity $\eta = 10^{-4}$ along the diagonal. In the simulation of the algorithm, we observe that the algorithm is not too sensitive to these choices.

From equations (31) and (36), we obtain the four limiting spectral moments $(1/n)\sum_{i=1}^{n} X_i (X_i)^\top$ of the sample covariance matrices as follows:

$$\begin{aligned} \beta_1 &= \mu_1, \\ \beta_2 &= y\mu_1^2 + \mu_2, \\ \beta_3 &= y^2\mu_1^3 + 3y\mu_1\mu_2 + \mu_3, \\ \beta_4 &= y^3\mu_1^4 + 3y^2\mu_1^2\mu_2 + 2y(\mu_2^2 + 2\mu_1\mu_3) + \mu_4. \end{aligned} \tag{51}$$

The authors of [23, 24] state that, as the number of moments increases (especially when $n \geq 5$), the maximum entropy method becomes sensitive to numerical differences between the order moments, the gradient, and the state of the Hessian matrix, as well as improper arithmetic accuracy becomes very sensitive and can be improved by introducing different basis functions, such as a drift with Chebyshev polynomial basis functions with drift to overcome these effects. In this section, to simplify the operations, we estimate the maximum entropy density function in terms of 4th-order spectral moments. The value of (d; n; a; b) determines the value of each spectral moment. Figures 1–6 show histograms of eigenvalues and maximum entropy spectral density curves (MaxEntPDF) for different values of (d; n; a; b). From these plots, it can be seen that the maximum entropy spectral density function curve can match histograms of the eigenvalues of the covariance matrices very well.

Input: $\beta = (\beta_0, \ldots, \beta_N)$, Ending accuracy ε, Diitter variance in Hessian $\eta = 10^{-4}$;
Output: Coefficients $\{\lambda_i\}$;
   initial $\lambda_0 = (-\ln(\sqrt{2\pi}), 0, \ldots, 0)$, $k = 0$;
  repeat
    Compute gradient $g = (g_0, \ldots, g_N)$: $g_j = \beta_j - \int_\Omega x^j \exp(\sum_{i=0}^N \lambda_{ki} x^i) dx$;
    Compute Hessian G;
    Direction of Newton's law: $d_k = -(G + \eta I)^{-1} \cdot g$;
    Next iteration point: $\lambda_{k+1} = \lambda_k + d_k$;
    $k = k+1$;
  until $\|\beta - \widetilde{\mu}\|_2 < \varepsilon$.

ALGORITHM 1: The proposed algorithm.
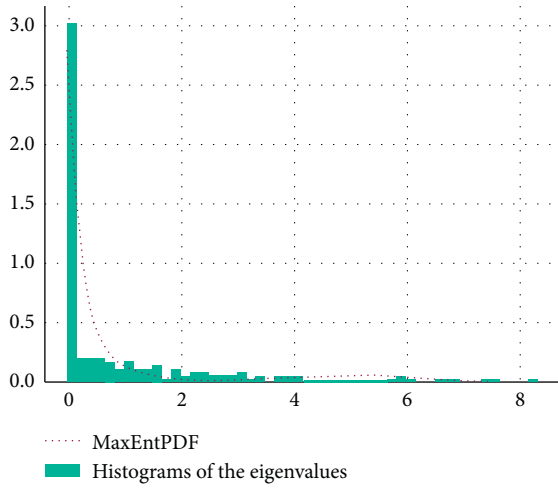
FIGURE 1: $d = 200$, $n = 100$, alpha = 0.5, and beta = 0.3.
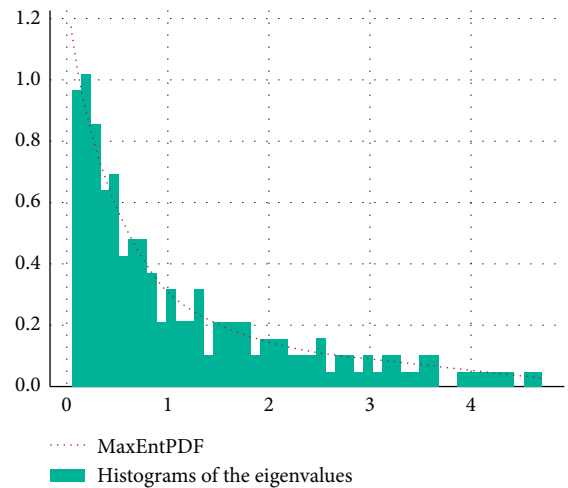
FIGURE 3: $d = 200$, $n = 400$, alpha = 0.5, and beta = 0.3.
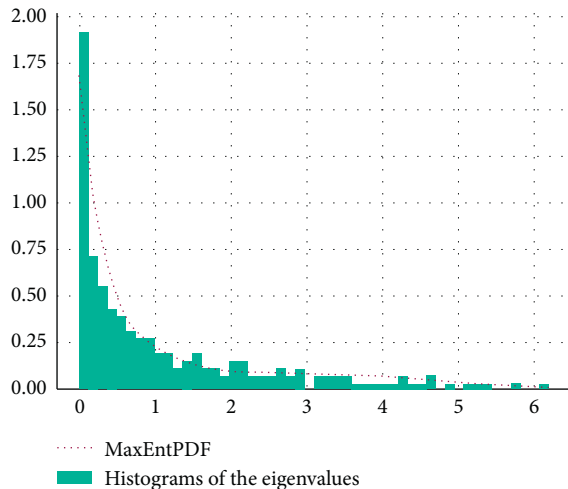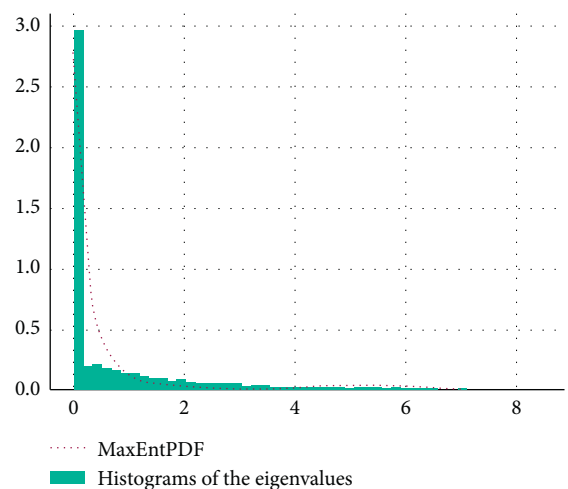

FIGURE 2: $d = 200$, $n = 200$, alpha = 0.5, and beta = 0.3.
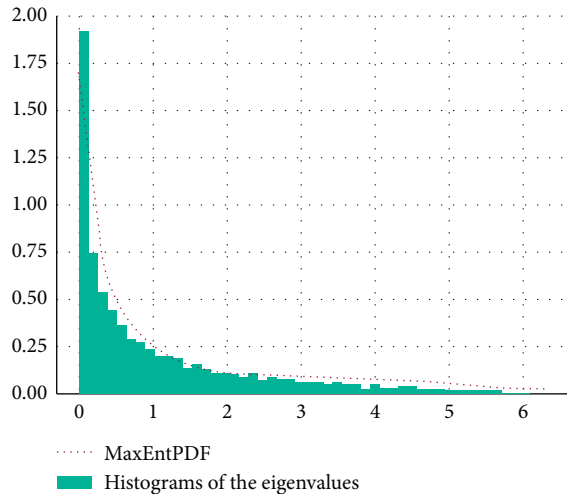

FIGURE 4: $d = 1000$, $n = 500$, alpha = 0.5, and beta = 0.3.

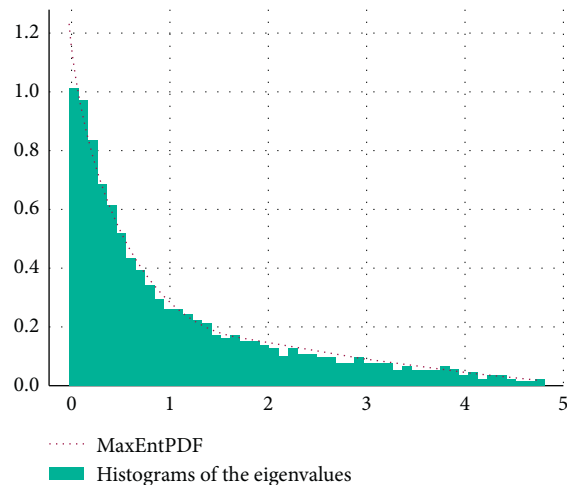Figure 5: $d = 1000$, $n = 1000$, alpha $= 0.5$, and beta $= 0.3$.

Figure 6: $d = 1000$, $n = 2000$, alpha $= 0.5$, and beta $= 0.3$.

4. Conclusion

In this paper, we investigate the LSD of large-dimensional sample covariance matrix generated by the first-order variable-coefficient PAR process. We consider the process be interrupted by certain factors that the odd terms and the even terms of the coefficient $\{\alpha_v\}$ converge to α and β, respectively, and that $|\alpha| \in (0, 1)$ and $|\beta| \in (0, 1)$. The main contributions in this paper include the following. (i) For the first-order variable-coefficient PAR model, the explicit form of limiting spectral moments for $\mathbf{Q}^{-1}(\mathbf{Q}^{-1})^{\top}$ is given; through verifying the Carleman condition, we derive the explicit spectral moments of $(1/n)\sum_{i=1}^{n}\mathbf{X}_i\mathbf{X}_i^{\top}$. (ii) A framework of the maximum entropy-based method is provided for estimate the LSD. We have some ideas for the next step of the research. On the one hand, when generalizing the convergence of the coefficients of the PAR model to a more general structure, how about obtaining an expression for the spectral moments of the limiting covariance

matrix? On the other hand, it is of importance to extend the method presented in this paper to PAR with larger order.

Data Availability

No data were used to support this study.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] P. H. Franses and R. Paap, "Periodic time series models," *Journal of the Royal Statal Society Series A (Stats in Society)*, vol. 168, no. 3, 2004.

[2] W. R. Bell, S. H. Holan, and T. S. McElroy, *Economic Time Series: Modeling and Seasonality*, CRC Press, Boca Raton, FL, USA, 2012.

[3] B. Aliat and F. Hamdi, "On Markov-switching periodic ARMA models," *Communications in Statistics-Theory and Methods*, vol. 47, no. 2, pp. 344–364, 2018.

[4] E. P. Wigner, "Characteristic vectors of bordered matrices with infinite dimensions," *Annals of Mathematics*, vol. 62, no. 3, pp. 548–564, 1955.

[5] E. P. Wigner, "On the distribution of the roots of certain symmetric matrices," *Annals of Mathematics*, vol. 2, no. 67, pp. 325–327, 1958.

[6] V. A. Marchenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Matematicheskii Sbornik*, vol. 114, no. 4, pp. 507–536, 1967.

[7] Y. Q. Yin and P. R. Krishnaiah, "A limit theorem for the eigenvalues of product of two random matrices," *Journal of Multivariate Analysis*, vol. 13, no. 4, pp. 489–507, 1983.

[8] Z. D. Bai, B. Miao, and B. Jin, "On limit theorem for the eigenvalues of product of two random matrices," *Journal of Multivariate Analysis*, vol. 98, no. 1, pp. 76–101, 2007.

[9] J. W. Silverstein, "Strong convergence of the empirical distribution of eigenvalues of large dimensional random matrices," *Journal of Multivariate Analysis*, vol. 55, no. 2, pp. 331–339, 1995.

[10] Y. Q. Yin, "Limiting spectral distribution for a class of random matrices," *Journal of Multivariate Analysis*, vol. 20, no. 1, pp. 50–68, 1986.

[11] Z. Bai and W. Zhou, "Large sample covariance matrices without independence structures in columns," *Statistica Sinica*, vol. 18, no. 2, pp. 425–442, 2008.

[12] B. Jin, C. Wang, B. Miao, and M.-N. Lo Huang, "Limiting spectral distribution of large-dimensional sample covariance matrices generated by VARMA," *Journal of Multivariate Analysis*, vol. 100, no. 9, pp. 2112–2125, 2009.

[13] Y. Zhang, J. Chen, B. Cui, and B. Tian, "Nonparametric estimate of spectral density functions of sample covariance matrices generated by VARMA models," *Communications in Statistics—Theory and Methods*, vol. 49, no. 1, pp. 1–10, 2020.

[14] J. Yao, "A note on a Marčenko-Pastur type theorem for time series," *Statistics & Probability Letters*, vol. 82, no. 1, pp. 22–28, 2012.

[15] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications*, University of California Press, Berkeley, CA, USA, 1958.

[16] R. T. Gregory, *A Collection of Matrices for Testing Computational Algorithms*, Wiley-Interscience, New York, NY, USA, 1969.

[17] F. Fathi, M. A. Fariborzi Araghi, and S. A. Shahzadeh Fazeli, "Two different inverse eigenvalue problems for nonsymmetric tridiagonal matrices," *Journal of Algorithms and Computation*, vol. 52, no. 2, pp. 137–148, 2020.

[18] D. A. Bini, L. Gemignani, and F. Tisseur, "The ehrlich--aberth method for the nonsymmetric tridiagonal eigenvalue problem," *SIAM Journal on Matrix Analysis and Applications*, vol. 27, no. 1, pp. 153–175, 2005.

[19] Z. D. Bai, "Methodologies in spectral analysis of large dimensional random matrices, a review," *Statistica Sinica*, vol. 9, pp. 611–677, 2008.

[20] L. R. Mead and N. Papanicolaou, "Maximum entropy in the problem of moments," *Journal of Mathematical Physics*, vol. 25, no. 8, pp. 2404–2417, 1984.

[21] D. Granziol, B. Ru, S. Zohren, X. Dong, M. Osborne, and S. Roberts, "MEME: an accurate maximum entropy method for efficient approximations in large-scale machine learning," *Entropy*, vol. 21, no. 6, 2019.

[22] L. Gao and D. Han, "Methods of moment and maximum entropy for solving nonlinear expectation," *Mathematics*, vol. 7, no. 1, 2019.

[23] K. Bandyopadhyay, A. K. Bhattacharya, P. Biswas, and D. A. Drabold, "Maximum entropy and the problem of moments: a stable algorithm," *Physical review. E, Statistical, nonlinear, and soft matter physics*, vol. 71, no. 5, pp. 057701–057751, 2005.

[24] H. Gotovac and B. Gotovac, "Maximum entropy algorithm with inexact upper entropy bound based on Fup basis functions with compact support," *Journal of Computational Physics*, vol. 228, no. 24, pp. 9079–9091, 2009.