Caspian Journal of Mathematical Sciences (CJMS)
University of Mazandaran, Iran
http://cjms.journals.umz.ac.ir
ISSN: 2676-7260

CJMS. **10**(2)(2021), 280-288

On the Empirical Spectral Distribution of Lag-Covariance Matrix in Singular Spectrum Analysis

Mahdi Kalantari [1] and Hormoz Rahmatan [2]
[1] Department of Statistics, Payame Noor University, 19395–4697, Tehran, Iran.
[2] Department of Mathematics, Payame Noor University, 19395–4697, Tehran, Iran.

Abstract. Singular Spectrum Analysis (SSA) is a non-parametric and rapidly developing method of time series analysis. Recently, this technique receives much attention in a wide variety of fields. In SSA, a special matrix, which is called lag-covariance matrix, plays a pivotal role in analyzing stationary time series. The objective of this paper is to examine whether the Empirical Spectral Distribution (ESD) of lag-covariance matrix converges to Marčenko–Pastur distribution or not. Such limiting distribution can help us to provide more reliable statistical inference when encountering with high-dimensional data. Moreover, a simulation study is performed and some tools of Random Matrix Theory (RMT) are used.

Keywords: Singular Spectrum Analysis, Random Matrix Theory, Empirical Spectral Distribution, Marčenko–Pastur Distribution, Lag-Covariance Matrix.

2000 Mathematics subject classification: 15B52, 60B20, 15A18; Secondary 37M10.

[2]Corresponding author: h_rahmatan@pnu.ac.ir
Received: 18 May 2020
Revised: 07 Februaty 2021
Accepted: 08 February 2021

1. Introduction

High-dimensional statistics is one of fields of statistics that studies data whose dimension (the number of features or variables) is larger than dimensions considered in classical multivariate statistics. The continued growth of large volume of more complex data sources obliges us to incorporate different mathematical tools into the statistical analysis. Random Matrix Theory (RMT) is a such mathematical tool that plays a pivotal role in modern high dimensional statistical inference [1]. It has many applications in statistics including hypothesis testing, clustering, regression analysis, Principal Component Analysis (PCA), Factor Analysis (FA), and Multivariate Analysis of Variance (MANOVA). A comprehensive review of RMT focusing on several application areas in statistics can be found in [2]. Other applications of RMT are in physics, biology, wireless communications, computer science, economics and finance [2]. One of important topics in RMT that plays a central role in studying the properties of the spectrum is the Empirical Spectral Distribution (ESD) of a random matrix. Describing the asymptotic convergence of the ESD to a proper probability distribution is of great interest in RMT. Key contributions in this framework are *semicircle law* and *Marčenko–Pastur law*, which are explained in the next section.

Singular Spectrum Analysis (SSA) is a non-parametric forecasting and filtering method that has many applications in a variety of fields such as signal processing, medicine, biology, genetics, engineering, finance, economics and time series analysis. For such examples of several applications of SSA see [3, 4, 5, 6, 7, 8, 9, 10, 11]. A whole and precise details on the theory and applications of SSA can be found in [12, 13, 14]. For a recent comprehensive review of SSA and description of its modifications and extensions, we refer the interested reader to [15]. In this paper, we focus on lag-covariance matrix in SSA because this matrix is at the core of analyzing stationary time series in SSA framework. The aim of this study is to examine whether the limiting distribution of ESD of lag-covariance matrix is Marčenko–Pastur distribution or not. In doing so, a simulation study is performed.

The remainder of this paper is organized as follows. Section 2 briefly presents a theoretical backgrounds and Section 3 is dedicated towards a simulation study. The conclusions and summary are presented in Section 4.

2. Theoretical Backgrounds

In this section, we present some key definitions and two theorems that play fundamental role in the RMT.

Definition 2.1. A random matrix is a random variable that takes its values in the space of matrices. In other words, it is just a matrix whose elements are random variables.

Definition 2.2. Suppose that $\mathbf{A}$ is an $n \times n$ matrix with eigenvalues $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$. The Empirical Spectral Distribution (ESD) or spectral measure of $\mathbf{A}$, which is denoted by $\mu(\mathbf{A})$, is the empirical distribution of its eigenvalues, namely

$$\mu(\mathbf{A}) = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i}, \tag{2.1}$$

where δ_y is the Dirac mass at y.

Note that μ puts equal mass on each eigenvalue of $\mathbf{A}$. When $\mathbf{A}$ is a random matrix, $\mu(\mathbf{A})$ is a random measure on $(\mathbb{R}, \mathscr{B})$ [16]. If $\mathbf{A}$ is Hermitian, the eigenvalues of $\mathbf{A}$ are real and consequently, the empirical distribution function of $\mathbf{A}$ can be defined as follows.

Definition 2.3. The empirical distribution function of $\mathbf{A}$, which is denoted by $F^{\mathbf{A}}(x)$, is defined as

$$F^{\mathbf{A}}(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{\{\lambda_i \leq x\}}, \qquad x \in \mathbb{R}, \tag{2.2}$$

where $\mathbf{1}_B$ is the indicator function of set B.

It is noteworthy that many statistics associated with a random matrix $\mathbf{A}$ can be expressed as a linear functional of its ESD, or a *linear spectral statistic*, that is, a function of the form $\int g(x) dF^{\mathbf{A}}(x)$ for some suitably regular function g [2]. For example,

$$\log(\det(\mathbf{A})) = \sum_{i=1}^{n} \log \lambda_i = n \int \log(x) dF^{\mathbf{A}}(x), \tag{2.3}$$

and the kth moment of the ESD of $\mathbf{A}$ equals $tr(\mathbf{A}^k)$, that is,

$$tr(\mathbf{A}^k) = n \int x^k dF^{\mathbf{A}}(x). \tag{2.4}$$

Therefore, knowing the asymptotic behavior of the ESD can help in studying the behavior of linear spectral statistics. In the seminal paper [17], Wigner proved that the spectral measure of a wide class of symmetric random matrices of dimension n converges to the semicircle law, as $n \to \infty$. Wigner matrix is defined as follows.

Definition 2.4. Wigner matrix is a square Hermitian matrix whose diagonal entries are independent and identically distributed (i.i.d) real random variables with mean 0 and variance 1, and those above the diagonal are i.i.d. complex random variables with mean 0 and variance 1.

There have been numerous further developments that determined in particular the necessary and sufficient conditions for the convergence of the ESD of the Wigner matrix. The following theorem states the result under the weakest moment conditions [2].

Theorem 2.5. *Suppose that* $\mathbf{A}$ *is a Wigner matrix. If* $n \to \infty$*, then the ESD of* $\mathbf{A}/\sqrt{n}$ *almost surely converges in distribution to the semicircle law with probability density function (p.d.f.) given by*

$$f(x) = \frac{1}{2\pi}\sqrt{4-x^2}\,\mathbf{1}_{[-2,2]}(x) \tag{2.5}$$

Marčenko and Pastur [18] derived the limiting distribution of the ESD of a sample covariance matrix, which is defined as $\mathbf{S} = \frac{1}{n}\mathbf{X}\mathbf{X}^*$, assuming that the fourth moments of the entries of the data $p \times n$ matrix $\mathbf{X}$ are finite. Since then, many researches have contributed to weakening the conditions on the matrix entries; see, for example [19, 20, 21]. The following theorem is under the minimal moment conditions [2].

Theorem 2.6. *Suppose that* $\mathbf{X}$ *is a* $p \times n$ *matrix with i.i.d. real-or complex-valued entries with mean 0 and variance 1. Suppose also that* $\lim_{n\to\infty} \frac{p}{n} = \gamma > 0$*. Then, as* $n \to \infty$*, the ESD of* $\mathbf{S} = \frac{1}{n}\mathbf{X}\mathbf{X}^*$ *converges almost surely in distribution to a nonrandom distribution, known as the Marčenko–Pastur law and denoted by* F_γ*. If* $\gamma \in (0,1]$*, then* F_γ *has the p.d.f.*

$$f_\gamma(x) = \frac{1}{2\pi\gamma x}\sqrt{(b_+(\gamma)-x)(x-b_-(\gamma))}\,\mathbf{1}_{[b_-(\gamma),b_+(\gamma)]}(x), \tag{2.6}$$

where $b_\pm(\gamma) = \left(1 \pm \sqrt{\gamma}\right)^2$*. If* $\gamma \in (1,\infty)$*, then* F_γ *is a mixture of a point mass 0 and the p.d.f.* $f_{1/\gamma}$ *with weights* $1 - 1/\gamma$ *and* $1/\gamma$*, respectively.*

Figure 1 shows the density function of the Marčenko–Pastur distribution for some values of γ.

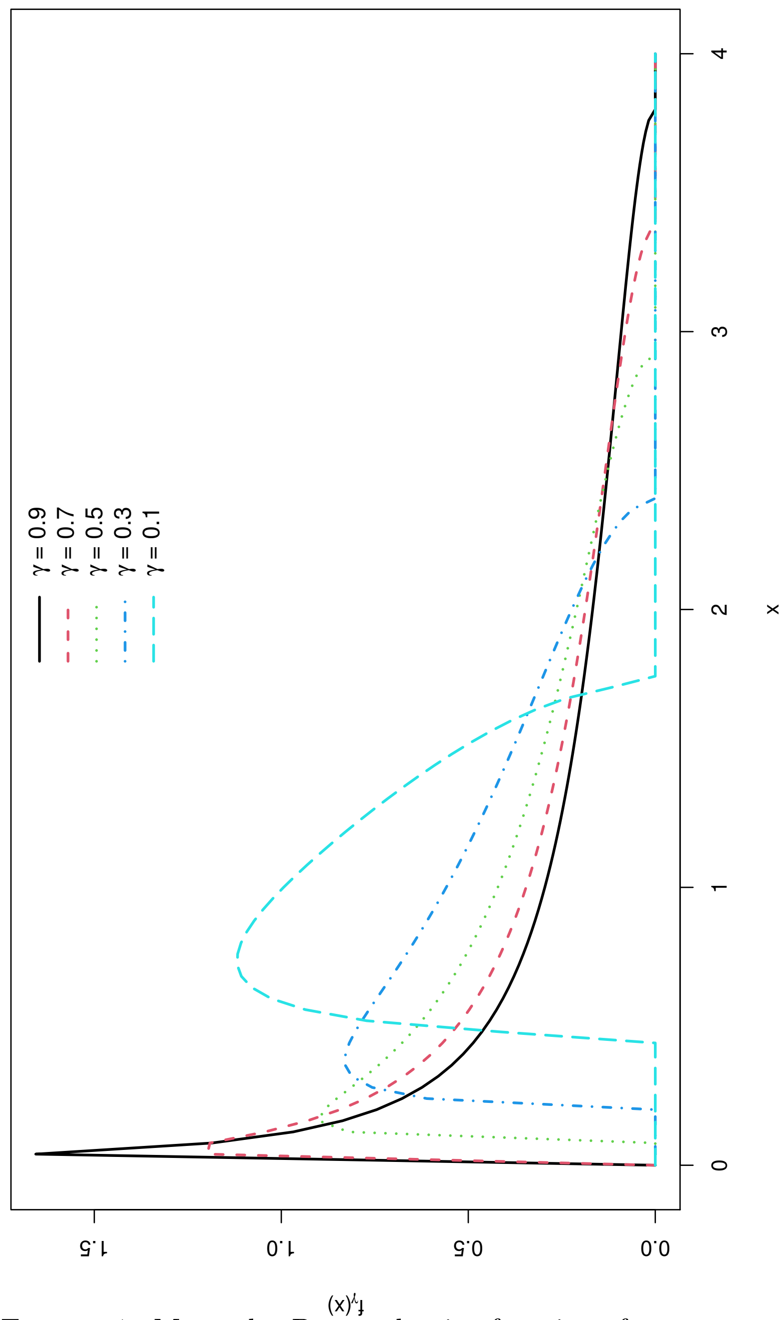

Figure 1. Marčenko–Pastur density functions for $\gamma = 0.1, 0.3, 0.5, 0.7, 0.9$.

The basic SSA technique consists of four steps: embedding, Singular Value Decomposition (SVD), grouping, and diagonal averaging. In the embedding step, the realization $y_1, \ldots, y_N$ of time series $\{Y_t\}_{t\geq 1}$ is transformed to the sub-series $X_1, \ldots, X_K$, where $X_i = (y_i, \ldots, y_{i+L-1})^T \in \mathbb{R}^L$ and $K = N-L+1$. The vectors X_i are called *L-lagged vectors.* The single choice of this step is the *Window Length* L, which is an integer such that $2 \leq L \leq N/2$. The output of the embedding step is the *trajectory* matrix $\mathbf{X}$, whose columns are the L-lagged vectors,

$$\mathbf{X} = [X_1 : \cdots : X_K] = \begin{pmatrix} y_1 & y_2 & y_3 & \cdots & y_K \\ y_2 & y_3 & y_4 & \cdots & y_{K+1} \\ y_3 & y_4 & y_5 & \cdots & y_{K+2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_L & y_{L+1} & y_{L+2} & \cdots & y_N \end{pmatrix}_{L\times K} \tag{2.7}$$

The trajectory matrix $\mathbf{X}$ is also a *Hankel* matrix in the sense that all elements on the anti-diagonals are equal. In SSA literature, the lag-covariance matrix is defined as $\mathbf{S} = \frac{1}{K}\mathbf{X}\mathbf{X}^T$. Note that since $K = N - L + 1$ and it is assumed that $L \leq N/2$, we have $\gamma = \frac{L}{K} \in (0, 1)$. In the next section, we try to check whether the limiting distribution of ESD of the lag-covariance matrix $\mathbf{S}$ is the Marčenko–Pastur distribution (2.6) or not.

3. Simulation Results

In order to perform a simulation study, first, we simulate N normally distributed random variable with zero mean and unit variance. Then the trajectory matrix $\mathbf{X}$ is constructed using L-lagged vectors. Here, the p-value of the Kolmogorov-Smirnov test is compared with the significance level of test (α), which is considered at three levels 1%, 5% and 10%, to assess fitting the Marčenko–Pastur distribution to the ESD of lag-covariance matrix $\mathbf{S}$. The simulation was repeated 1000 times for each of different combinations of (L, K) and finally, the percent of times that the Marčenko–Pastur distribution fitted was calculated. We implemented the `RMTstat` package of free-available `R` software to use the Marčenko–Pastur distribution. More details on this package can be found in [22].

Table 1. Percent of fitting Marčenko–Pastur distribution

$\gamma = \frac{L}{K}$	α	(L, K)				
		(50, 500)	(100, 1000)	(150, 1500)	(200, 2000)	(300, 3000)
0.1	1%	97.8	95.8	95.7	92.5	90.9
	5%	90.1	87.3	85.7	79.8	71.7
	10%	81.8	77.6	73.8	67.2	58.9
		(150, 500)	(300, 1000)	(450, 1500)	(600, 2000)	(900, 3000)
0.3	1%	95.3	90.9	81.2	72.8	50
	5%	84.5	73.5	54.9	37.3	17.6
	10%	71.2	58	39.3	21.8	7.9
		(250, 500)	(500, 1000)	(750, 1500)	(1000, 2000)	(1500, 3000)
0.5	1%	91.9	77.5	54.3	30.6	4.9
	5%	75	42.8	20.9	5.6	0.2
	10%	59.6	24.9	8.6	1.3	0
		(350, 500)	(700, 1000)	(1050, 1500)	(1400, 2000)	(2100, 3000)
0.7	1%	87.6	52.8	17.2	2.8	0.1
	5%	61.7	15.4	1.1	0	0
	10%	43.6	6.2	0.3	0	0
		(450, 500)	(900, 1000)	(1350, 1500)	(1800, 2000)	(2700, 3000)
0.9	1%	84.3	35.6	4.2	0.3	0
	5%	52.4	5.5	0	0.1	0
	10%	29.3	1.3	0	0	0

The percent of fitting Marčenko–Pastur distribution is reported in Table 1. As can be seen from this table, for each γ and at each level of significance (α), the percent of fitting decreases as L and K increase. Therefore, it can be concluded that the limiting distribution of ESD tends to be far from the Marčenko–Pastur distribution, as $L, K \to \infty$. In addition, for a fixed γ, the greatest fitting percent corresponds to the smallest (L, K). Also, it can be easily seen that at each level of α, the largest fitting percent, which corresponds to $K = 500$, falls down as γ rises up. On the other words, the best fitting percent is achieved when $\gamma = 0.1$, $L = 50$, and $K = 500$. In summary, it seems from the results of Table 1 that the Marčenko–Pastur distribution is not the limiting distribution of ESD of the lag-covariance matrix $\mathbf{S}$, especially for larger values of γ.

4. Conclusion

In this paper, we have performed a simulation study to check whether the Marčenko–Pastur distribution can be used as a limiting distribution of ESD of lag-covariance matrix $\mathbf{S}$ or not. The results of the present study, which is based on the p-values of the Kolmogorov-Smirnov test, have shown that the Marčenko–Pastur distribution can not be applied as a limiting distribution of ESD of lag-covariance matrix $\mathbf{S}$ in SSA. A precise look at the trajectory matrix **X** in (2.7) reveals that the elements on the anti-diagonals are not independent, because the matrix **X** is a Hankel matrix and hence, all elements on the anti-diagonals are equal. However, it is assumed in Theorem 2.6 that matrix entries should be independent. The results of our simulation provides sound evidence that the Marčenko–Pastur distribution is not a limiting distribution of ESD of lag-covariance matrix **S** in SSA and finding such limiting distribution needs further investigation.

References

[1] J. Yao, S. Zheng and Z. Bai, Large sample covariance matrices and high dimensional data analysis, Cambridge University Press, New York, 2015.

[2] D. Paul and A. Aue, Random matrix theory in statistics: A review, *Journal of Statistical Planning and Inference*, **150**(2014), 1–29.

[3] M. Atikur Rahman Khan and D. S. Poskitt, Forecasting stochastic processes using singular spectrum analysis: Aspects of the theory and application, *International Journal of Forecasting*, **33**(2017), 199–213.

[4] J. Arteche and J. Garcia-Enriquez, Singular Spectrum Analysis for signal extraction in Stochastic Volatility models, *Econometrics and Statistics*, **1**(2017), 85–98.

[5] H. Hassani, M. R. Yeganegi, and E. S. Silva, A New Signal Processing Approach for Discrimination of EEG Recordings, *Stats*, **1**(2018), 155–168.

[6] S. M. M. Safi, M. Mohammad Pooyan, and A. M. Nasrabadi, Improving the performance of the SSVEP-based BCI system using optimized singular spectrum analysis (OSSA), *Biomedical Signal Processing and Control*, **46**(2018), 46–58.

[7] R. Mahmoudvand and P. C. Rodrigues, Predicting the Brexit Outcome Using Singular Spectrum Analysis, *Journal of Computational Statistics and Modeling*, **1 (1)** (2018), 9–15.

[8] S. Lahmiri, Minute-ahead stock price forecasting based on singular spectrum analysis and support vector regression, *Applied Mathematics and Computation*, **320**(2018), 444–451.

[9] A. Saayman and J. Klerk, Forecasting tourist arrivals using multivariate singular spectrum analysis, *Tourism Economics*, **25 (3)** (2019), 330–354.

[10] H. Hassani, A. Rua, E. S. Silva, and D. Thomakos, Monthly forecasting of GDP with mixed-frequency multivariate singular spectrum analysis, *International Journal of Forecasting*, **35**(2019), 1263–1272.

[11] D. S. Poskitt, On Singular Spectrum Analysis and Stepwise Time Series Reconstruction. *J. Time Ser. Anal.* (2019), DOI: 10.1111/jtsa.12479.

[12] N. Golyandina, V. Nekrutkin, and A. Zhigljavsky, Analysis of Time Series Structure: SSA and related techniques, Chapman & Hall/CRC, London, 2001.

[13] N. Golyandina and A. Zhigljavsky, Singular Spectrum Analysis for Time Series, Springer Briefs in Statistics, Springer, 2013.

[14] N. Golyandina, A. Korobeynikov, and A. Zhigljavsky, Singular Spectrum Analysis with R, Springer-Verlag, Berlin Heidelberg, 2018.

[15] N. Golyandina, Particularities and commonalities of singular spectrum analysis as a method of time series analysis and signal processing, *WIREs Computational Statistics*, (2020), e1487.

[16] W. Bryc, A. Dembo, and T. Jiang, Spectral Measure of Large Random Hankel, Markov and Toeplitz Matrices, *The Annals of Probability*, **34 (1)** (2006), 1–38.

[17] E. P. Wigner, On the distribution of the roots of certain symmetric matrices, *Ann. of Math.*, **67 (2)** (1958), 325–327.

[18] V. Marčenko, and L. Pastur, Distribution of eigenvalues for some sets of random matrices, *Mathematics of the USSR-Sbornik*, **1**(1967), 457–483.

[19] K. W. Wachter, The strong limits of random matrix spectra for sample matrices of independent elements, *Annals of Probability*, **6** (1978), 1–18.

[20] Y. Q. Yin, Limiting spectral distribution for a class of random matrices, *Journal of Multivariate Analysis*, **20**(1986), 50–68.

[21] Y. Q. Yin, Z. D. Bai, and P. R. Krishnaiah, Limiting behavior of the eigenvalues of a multivariate F matrix, *Journal of Multivariate Analysis*, **13** (1983), 508–516.

[22] I. M. Johnstone, Z. Ma, P. O. Perry, and M. Shahram, RMTstat: Distributions, Statistics and Tests derived from Random Matrix Theory, R package version 0.3 (2014).