University of Windsor
Scholarship at UWindsor

Major Papers | Theses, Dissertations, and Major Papers

June 2024

On Limiting Distributions for Eigenvalue Spectra of Sample Correlation Matrices from Heavy-Tailed Populations: Literature Review

Sachini Sandhareka Wijesundara
University of Windsor, wijesuns@uwindsor.ca

Follow this and additional works at: https://scholar.uwindsor.ca/major-papers

Part of the Probability Commons, Statistical Methodology Commons, Statistical Models Commons, and the Statistical Theory Commons

Recommended Citation

Wijesundara, Sachini Sandhareka, "On Limiting Distributions for Eigenvalue Spectra of Sample Correlation Matrices from Heavy-Tailed Populations: Literature Review" (2024). *Major Papers*. 308.
https://scholar.uwindsor.ca/major-papers/308

This Major Research Paper is brought to you for free and open access by the Theses, Dissertations, and Major Papers at Scholarship at UWindsor. It has been accepted for inclusion in Major Papers by an authorized administrator of Scholarship at UWindsor. For more information, please contact scholarship@uwindsor.ca.

On Limiting Distributions for Eigenvalue Spectra of Sample Correlation Matrices from Heavy-Tailed Populations: Literature Review

By

Sachini Wijesundara

A Major Research Paper
Submitted to the Faculty of Graduate Studies
through the Department of Mathematics and Statistics
in Partial Fulfillment of the Requirements for
the Degree of Master of Science
at the University of Windsor

Windsor, Ontario, Canada

2024

©2024 Sachini Wijesundara

On Limiting Distributions for Eigenvalue Spectra of Sample Correlation Matrices from Heavy-Tailed Populations: Literature Review

by

Sachini Wijesundara

APPROVED BY:

M. Belalia
Department of Mathematics and Statistics

S. Nkurunziza, Advisor
Department of Mathematics and Statistics

May 28, 2024

Author's Declaration of Originality

I hereby certify that I am the sole author of this major paper and that no part of this major paper has been published or submitted for publication.

I certify that, to the best of my knowledge, my major paper does not infringe upon anyone's copyright nor violate any proprietary rights and that any ideas, techniques, quotations, or any other material from the work of other people included in my major paper, published or otherwise, are fully acknowledged in accordance with the standard referencing practices. Furthermore, to the extent that I have included copyrighted material that surpasses the bounds of fair dealing within the meaning of the Canada Copyright Act, I certify that I have obtained a written permission from the copyright owner(s) to include such material(s) in my major paper and have included copies of such copyright clearances to my appendix.

I declare that this is a true copy of my major paper, including any final revisions, as approved by my thesis committee and the Graduate Studies office, and that this major paper has not been submitted for a higher degree to any other University or Institution.

Abstract

This major paper offers an extensive review of literature concerning the limiting distributions for the eigenvalue spectrum of sample correlation matrices from a p-dimensional population, where both the dimension p and the sample size n grow to infinity. The study systematically categorizes the reviewed literature based on underlying assumptions regarding the data characteristics. Specifically, it examines several distinct cases: the independent and identically distributed (i.i.d) case with finite fourth moments, the i.i.d case with infinite fourth moments, the i.i.d case with infinite second moments, and scenarios where rows and columns of the data are linearly dependent. Additionally, the major paper provides brief insights into the methodologies employed in the reviewed papers, offering a glimpse into the diverse analytical techniques utilized to investigate the limiting distributions of eigenvalues in correlation matrices.

Dedication

This dedication is to all those who have supported and encouraged me throughout my journey.

To my husband, whose unwavering support and strength have been my backbone throughout this journey. His presence and encouragement have been invaluable to me, and I am deeply grateful for his love and dedication.

To my dad who has taught me a lot in my life and study - your knowledge, wisdom, and experience have been invaluable in shaping who I am today. And to all those who have touched my life in some way - thank you for being a part of my journey and for helping me become the person I am today.

To my classmates who challenged me - thank you for pushing me out of my comfort zone and helping me grow.

Acknowledgements

I would like to sincerely express my most profound gratitude towards my supervisor Prof. Sévérien Nkurunziza, whose input helped me immensely. With his input, I was able to look at my research with a different perspective and a more critical eye.

I would also like to thank Dr. Belalia for being my department reader and for the beneficial advice and suggestions for my major paper.

In addition, I humbly extend my thanks to the Department of Mathematics and Statistics and all concerned people who helped me in this regard.

Contents

List of Tables

List of Figures

List of Abbreviations

PCA	Principal Component Analysis
RMT	Random Matrix Theory
QM	Quantum Mechanics
ESD	Empirical Spectral Distribution
LSD	Limiting Spectral Distribution
CLT	Central Limit Theorem
LSSs	Linear Spectral Statistics
MP	Marčenko-Pastur
MCT	Moment Convergence Theorem

Chapter 1

Introduction

In the realm of statistical analysis, there has been a growing interest in studying heavy-tailed populations due to their widespread occurrence in various domains including finance, natural sciences, and telecommunications, as mentioned in Ahmad et al. (2020), LePage (2011), and Resnick (1997). According to LePage (2011), a distribution is classified as heavy-tailed when it exhibits a heavy-tailed pattern in at least one direction. Classical distributions cannot effectively manage data sets characterized by a heavy tail due to their basic lack of flexibility. For example, Pareto distribution is commonly used to describe financial data sets, but it may not suit all applications as stated in Guillen et al. (2011), while the Weibull model better fits the small loss behavior shown by Bhati and Ravi (2018). Ahmad et al. (2020) proposed that in such cases of describing financial data sets, employing heavy-tailed distributions as the candidate model is reliable and accurate. LePage (2011) defines a heavy right-tailed probability distribution function F on the real line as the ratio $\frac{1-F(x)}{e^{-tx}}$ having an infinite limit as x goes to infinity for each $t > 0$. The right-tail probabilities $1 - F(x)$ for this type of F exhibit a slower rate of decay to zero compared to exponential distributions. Furthermore, LePage (2011) observed that the probability distribution presents a heavy left-tailed trend when the probability function $F(-x)$ is heavy right-tailed.

Li et al. (2023), Parvin (2004), and Yao et al. (2015b) described that the sample covariance and sample correlation matrix are crucial in multivariate statistical analysis, while the population covariance matrix is utilized for hypothesis testing, canonical correlation analysis (CCA), and principal component analysis (PCA) as shown in

Yang et al. (2019), Bilodeau and Duchesne (2002), Johnstone (2001), and Anderson (1963). In large sample theory, it is considered that the classical limiting theorems in statistics are fully defined. This implies that the data dimension p remains constant as the sample size n approaches infinity, as stated by Morales-Jimenez et al. (2021). The claim being made aligns with the findings of Yao et al. (2015b), who asserts that the majority of asymptotic outcomes for sample covariance and correlation matrices are obtained based on this particular assumption. In a low-dimensional setting, Waternaux (1976) observed that the limiting joint distributions of the eigenvalues of the sample covariance matrix are multivariate normal when the population roots are simple and the underlying distribution is nonnormal. As mentioned in Davis (1977), the likelihood ratio test in limiting distribution for multivariate normal populations is equivalent to a normal quadratic form when the underlying distribution is not multivariate normal and the roots have multiplicity. The sample correlation matrix is effective with a small number of variables p and a large sample size n, and its statistical characteristics were widely recognized by Konishi (1979), Kollo and Neudecker (1993), and Boik (2003). Since the analysis of large-dimensional data presents a new problem, it is evident that most practitioners have adhered to the above idea until recently.

As indicated by Bai and Silverstein (2010), with the rapid growth and widespread use of computer science, numerous well-known approaches in multivariate analysis become inefficient or even deceptive when the data dimension p is not as small, say several tens. Therefore, recent applications frequently demonstrate large dimensionality, where p and n are of comparable magnitude. For example, Heiny and Yao (2022) considered the asymptotic regime (C_γ) defined as follows:

$$p = p_n \to \infty \quad \text{and} \quad \frac{p}{n} \to \gamma \in (0, \infty), \quad \text{as} \quad n \to \infty. \tag{C_γ}$$

Several recent studies in high-dimensional settings focus on examining the characteristics of the leading eigenvalues and their corresponding eigenvectors of the sample covariance and correlation matrices as described in Johnstone (2001), Jolliffe and

Cadima (2016), Hoyle and Rattray (2004), Baik and Silverstein (2006), and Lindsey (2004). In this regard, random matrix theory (RMT) has developed useful tools, where RMT traces back to the development of quantum mechanics (QM) in the 1940s and early 1950s. In QM, the energy levels of a system are described by the eigenvalues of a Hermitian operator A on a Hilbert space called the Hamiltonian, as mentioned in Bai and Silverstein (2010), and Couillet and Debbah (2011). Since the late 1950s, research on the limiting spectral analysis of large dimensional random matrices has attracted considerable interest among mathematicians, probabilists, and statisticians, as indicated in Bai and Silverstein (2010), and Yao et al. (2015b). According to Wigner (1958), a large-dimensional Wigner matrix's predicted spectral distribution follows the semicircular law. Further research by Arnold (1967), and Arnold (1971) studies the distribution of eigenvalues in large symmetric matrices with independent and identically distributed random variables, generalizing the large-dimensional Wigner matrix. Bai and Yin (1988a) demonstrated that the spectral distribution of a sample covariance matrix, when sufficiently normalized, tends to follow the semicircular law when the dimension is relatively less than the sample size.

Marchenko and Pastur (1967), and Pastur (1972) studied the distribution of eigenvalues for random matrices, focusing on two sets of random Hermitian matrices and one set of random unitary matrices. Numerous academics have established the asymptotic theory of spectrum analysis of large-dimensional sample covariance matrices. Bai and Yin (1988a) proved that as n tends to infinity, the largest eigenvalue $\lambda_{\max}(A)$ (same as $\lambda_1(A)$) of sample covariance matrices converges almost surely (a.s.) to a finite constant a. For the derivation, they have taken into account $A = (X_{ij}; 1 < i, j < \infty)$ as a symmetric infinite matrix. Subsequently, when n tends to infinity, Tikhomirov (2015) proved that the smallest eigenvalue $\lambda_m(A)$, where $a_{ij}; (i \leq 1, j < \infty)$ are i.i.d. real-valued random variables with zero mean and unit variance, converges a.s. to the edge of the M-P law. The study by Bai and Silverstein (2008) demonstrated that the linear spectral statistics converge a.s. to a nonrandom quantity due to the limiting behavior of the empirical spectral distribution. Furthermore, Bai and Silverstein (2008) shows that if the fourth moment is finite ($\mathbb{E}|X_{11}|^4 = 2$), they tend to have Gaussian limits. Pillai and Yin (2014) derived the edge universality (towards the Tracy-Widom law) of covariance matrices with independent real-valued entries, which satisfies the conditions $\mathbb{E}[X_{ij}] = 0$ and $\mathbb{E}[X_{ij}^2] = \frac{1}{M} < \infty$. The use of Gaussian covariance/correlation matrices in statistical problems is essential because it enables us to derive exact asymptotic distributions of test statistics without the need for matrix entries dependent on limiting distributional assumptions, as shown by Pillai and Yin (2012). Heiny and Yao (2022) points out that the absence of the finite fourth moment, where $\mathbb{E}[X^4] = \infty$, which is relevant in the case with light tails, the convergence theory for the eigenvalues and eigenvectors of sample covariance matrices is substantially different from the classical Marčenko and Pastur theory.

In the typical extreme value theory for i.i.d. random variables, as proven by Auffinger et al. (2009), the behavior of the largest matrix entries determines the asymptotic behavior of the top eigenvalues. This relies on regularly varying i.i.d entries, $a_{ij}; (1 \leq \mathrm{i} \leq n, 1 \leq \mathrm{j} \leq n)$ with $\alpha \in (0, 4)$ and a slowly varying function L at infinity defined as in (1.1). Also, Auffinger et al. (2009) proved that the largest eigenvalue converges to a Fréchet limit distribution.

$$P(|a_{ij}| > x) = L(x)x^{-\alpha}, \quad x > 0, \tag{1.1}$$

Similar methods for developing Wigner matrices with heavy-tailed entries may be discovered in the literature by Auffinger et al. (2009), Basrak et al. (2021), and Davis et al. (2016), which carefully examined the infinite fourth-moment case. A considerable amount of literature can be found in the case where the limiting distributions of sample covariance matrices were analyzed with infinite variance $\mathbb{E}[X^2] = \infty$. Assuming regular variation with index $\alpha \in (0, 2)$, it is demonstrated that the empirical spectral distribution (ESD) of the suitably normalized sample covariance matrix converges under the condition (C_γ). The paper by Belinschi et al. (2009), and Arous and Guionnet (2008) showed that the ESD converges weakly to a probability measure with infinite support. The specific form of this measure depends on the parameters α and γ.

Theoretical conclusions for high-dimensional sample correlation matrices primarily focus on the "null model" scenario, where data is independent. Jiang (2004) proved that the correlation matrices from i.i.d data with zero mean and unit variance show asymptotic similarities, indicating convergence of the empirical spectral distribution (F_A) to the Marčenko–Pastur distribution.

The study by Xiao and Zhou (2010), and Jiang (2004), assuming independent and identically distributed X_{ij}'s, found that the smallest and largest eigenvalues of R_n converge almost surely to the edge of the Markov-Pastur distribution. Gao et al. (2017) introduced a novel central limit theorem for the linear spectral statistic of high-dimensional sample correlation matrices. The theorem focused on scenarios where the dimensionality p and the sample size n are comparable. Similarly, an analogous result is derived for the described scenario. Bao et al. (2012), and Pillai and Yin (2012) showed that when independent, identically distributed entries exhibit subexponential decay in tail behavior, the sample correlation matrix has edge universality toward the Tracy-Widom law.

Despite considerable literature and applications, the theoretical characteristics of sample correlation, in particular the eigenvalue spectrum in large dimensions, are not well understood and are more complex. One primary explanation is that the entries of the standardized matrix lose their independence within each row, contrasting with the original data matrix. While the i.i.d of different rows remains unchanged, the dependency structure within rows changes significantly as in Heiny and Yao (2022). Due to several concerns, investigating limiting distributions for eigenvalues of sample correlation matrices is a critical task in heavy-tailed data. This research review explores the complex interactions between heavy-tailed populations and their sample correlation matrices, focusing on limiting distributions for eigenvalues.

The subsequent chapters of this major paper are organized as follows, Chapter 2 discusses the methods and algorithms employed for the spectral analysis of heavy-tailed data. Chapter 3 presents a comprehensive analysis of the literature findings, based on assumptions. The major paper ends with a conclusion in Chapter 4.

Chapter 2

Methodology

This chapter provides an overview of the various methods employed in the derivation of limiting distributions for sample correlation matrices based on heavy-tailed distributions. The organization of the methodology consists of a brief overview of random matrix theory and its relevance to the analysis of correlation matrices. Furthermore, it reviews the key concepts and results that are important in determining the limiting distributions of eigenvalues and understanding their characteristics.

Pillai and Yin (2012) introduced the standardized matrix, normalized by the Euclidean norm, and used Green functions to overcome dependence in correlation matrix entries. The moments of standardized matrix entries with raw data matrix moments were compared using the Stieltjes transform of the Empirical Spectral Distribution (ESD). Morales-Jimenez et al. (2021) built on the set of random matrix tools for studying spiked correlation models. Key techniques included the method of moments and the concentration of measure inequalities. Johnstone and Yang (2018) further established that these tools allowed for rigorous analysis of the asymptotic behavior of the eigenvalues in the spiked correlation models. The study by El Karoui (2009) utilizes various concentration measures in understanding the behavior of the limiting spectral distributions of a number of random matrix models. The same phenomenon was employed to deduce the spectral properties of sample correlation matrices from the corresponding properties of sample covariance matrices. The employed methodology includes truncation, centralization, and Stieltjes transform. Jiang (2004) used moment methods and random matrix tools to find limiting distributions of eigenvalues of R_n, while Heiny and Mikosch (2018), and Heiny and Yao (2022) used methods like moments, path-shortening algorithm, and graph counting combinatorics.

Non-parametric correlation matrices, such as Kendall's τ and Spearman's ρ, have recently acquired significance for analyzing heavy-tailed data. Li et al. (2023) conducted the study on the spectral properties of Kendall's rank correlation matrix. The Hoeffding decomposition approach was used to analyze the leading terms of the sign vector A_{ij} produced from the p-dimensional data sample $x_1, ..., x_n \in \mathbb{R}^p$. Additionally, the study utilizes the Stieltjes transform of the limiting spectral distribution of Kendall's rank correlation matrix. Further explanation of the methods described in this chapter can be found in the following references, Jiang (2004), Heiny and Yao (2022), Bai and Silverstein (2010), and Yao et al. (2015b).

The methodology is divided into the following sections. Section 2.1 describes the properties of the sample correlation matrix derived from heavy-tailed data. Also, it covers the methodology for computing the sample covariance and correlation matrices, including the standardization process, convergence properties, and the application of spectral analysis techniques to understand the behavior of eigenvalues. Section 2.2 includes the methodologies used to analyze the spectral properties of large-dimensional random matrices derived from heavy-tailed populations. This discussion encompasses techniques such as the moment method, stieltjes transform, and orthogonal polynomial decomposition, each offering insights into the behavior of eigenvalues. Additionally, this section describes key concepts such as the Marčenko-Pastur law, the path-shortening algorithm PS(I), and fundamental terminology from graph theory and combinatorics. Finally, in Section 2.3, we discuss the importance of the utilization of simulations to validate theoretical methods for estimating the limiting distributions of eigenvalues of sample correlation matrices.

2.1 Development of sample correlation matrix

The sample correlation matrix is derived by focusing the data from a heavy-tailed population. Consider a p-dimensional population $\mathbf{X}$ denoted as $\mathbf{X} = (X_1, \ldots, X_p) \in \mathbb{R}^p$ where the coordinates X_i are independent, non-degenerate random variables, which are identically distributed as a centered random variable ξ. The sample covariance matrix S and correlation matrix R can be constructed using the $p \times n$ data matrix $X = X_n = (X_{ij})_{1\le i\le p;1\le j\le n}$ given as;

$$S = S_n = \frac{1}{n} X X^{\top}, \tag{2.1}$$

$$R = R_n = \{\mathrm{diag}(S_n)\}^{-1/2} \quad S_n \quad \{\mathrm{diag}(S_n)\}^{-1/2} = YY^{\top}, \tag{2.2}$$

where $Y = Y_n = (Y_{ij})_{1\le i\le p;1\le j\le n}$ corresponds to the standardized matrix for the sample correlation matrix with the entries,

$$Y_{ij} = Y_{ij}^{(n)} = \frac{X_{ij}}{\sqrt{X_{i1}^2 + X_{i2}^2 + \cdots + X_{in}^2}}, \tag{2.3}$$

which depend on n. Throughout the paper, we often suppress the dependence on n in our notation. The corresponding eigenvalues of sample correlation matrix (R) and sample covariance matrix (S) are $\lambda_p(R) \le \cdots \le \lambda_2(R) \le \lambda_1(R)$ and $\lambda_p(S) \le \cdots \le \lambda_2(S) \le \lambda_1(S)$, respectively.

Recent advancements in sample correlation matrix R often involve the convergence of the normalizing denominator S_{ii} of Y_{ij}, as defined in relation (2.3) as $S_{ii} = \{X_{i1}^2 + X_{i2}^2 + \cdots + X_{in}^2\}/n$. The law of large numbers states that under the finite second-moment condition $\theta = \mathbb{E}[\xi^2] < \infty$, S_{ii} converges almost surely to θ as n tends to ∞ (Heiny and Yao, 2022). The analysis reveals a consistent trend across various studies in the field, highlighting the convergence phenomenon as a fundamental aspect. According to Lemma 2 in Bai and Yin (2008), uniform convergence of $\max_{1\le i\le p} |S_{ii} - \theta| \xrightarrow[n\to\infty]{a.s} 0$ is equivalent to the finite fourth moment condition $\mathbb{E}[\xi^4] < \infty$.

The eigenvalue perturbation inequality proposed by Weyl is crucial in understanding eigenvalue behavior in matrices undergoing perturbations. According to Weyl's inequality, the spectral properties of two Hermitian matrices, R and S, display asymptotic equivalence. That is, $\max_{1\le i\le p} |\lambda_i(R) - \theta^{-1}\lambda_i(S)| \xrightarrow[n\to\infty]{a.s} 0$ (El Karoui, 2009). This approach was mainly utilized in Johnstone (2001), Jiang (2004) and Bai and Silverstein (2010) for deriving that

$$\lambda_1(R) \xrightarrow[n\to\infty]{a.s} (1+\sqrt{\gamma})^2, \quad \text{and} \quad \lambda_p(R) \xrightarrow[n\to\infty]{a.s} (1-\sqrt{\gamma})^2,$$

and

$$\theta^{-1}\lambda_1(S) \xrightarrow[n\to\infty]{a.s} (1+\sqrt{\gamma})^2, \quad \text{and} \quad \theta^{-1}\lambda_p(S) \xrightarrow[n\to\infty]{a.s} (1-\sqrt{\gamma})^2,$$

with the growth condition, $\lim_{n\to\infty} \frac{p_n}{n} = \gamma \in (0, 1]$.

The concept of asymptotic equivalence of the spectral properties of R and S has been expanded to include those with a uniformly bounded spectrum and $\mathbb{E}[\xi^4(\log \xi)^{2+\epsilon}] < \infty$ as outlined in Theorem 1 of El Karoui (2009). The accuracy of estimating the error associated with the approximation is a crucial aspect of studies concerning the correlation matrix R. For example, Jiang (2004), El Karoui (2009), and Pillai and Yin (2012) show that this approximation error is negligible and the outcomes align for R and S. Heiny and Yao (2022) reference explores the infinite variance scenario where $\theta = \mathbb{E}[\xi^2] = \infty$. This approach reveals that the approximation argument loses validity, and the convergence limits of R no longer correspond to those of S.

Extensive studies on the sample correlation matrix R have mostly focused on the finite fourth moment assumption. In the case of infinite variance, the limiting spectral distribution of the sample covariance matrix S has been thoroughly studied by Belinschi et al. (2009), and Arous and Guionnet (2008). However, Heiny and Yao (2022) mentions that the literature provides minimal insight into the sample correlation matrix R in these settings.

Morales-Jimenez et al. (2021) examined a specific category of correlation matrix models, spiked models within the spectrum of sample correlation matrices assume that a few large or small eigenvalues of the population correlation matrix are distinctly separated from the remaining ones. Using random matrix theory, asymptotic first-order and distributional results for the principal eigenvalues and eigenvectors of sample correlation matrices were obtained by Johnstone (2001), and Morales-Jimenez et al. (2021). Paul (2007) derived theories about sample covariance matrices in the specific scenario of Gaussian data. Additionally, Morales-Jimenez et al. (2021) expanded upon these theories to include sample correlation matrices and extended them to non-Gaussian data.

2.2 Spectral analysis of large dimensional random matrices

According to Bai and Silverstein (2010), the limiting distributions for eigenvalues have been analyzed using theoretical and empirical approaches. These theoretical analyses involve deriving asymptotic results based on random matrix theory and related mathematical frameworks like graph theory concepts and combinatorics, as studied by Bai and Yin (1988a), Jiang (2004), and Heiny and Mikosch (2018). Empirical analyses are involved in simulating the eigenvalue distributions using computational methods such as Monte Carlo simulations, as stated by Adhikari and Friswell (2007).

Suppose that A is an $m \times m$ matrix with eigenvalues $\lambda_j, j = 1, 2, \ldots, m$, where $\lambda_m(A) \leq \ldots \leq \lambda_2(A) \leq \lambda_1(A)$. If all these eigenvalues are real (e.g., if A is Hermitian(see Appendix A 2)), we can define a one-dimensional distribution function called the empirical spectral distribution (ESD) of the matrix A as follows,

$$F_A = \frac{1}{m}\#\{j \leq m : \lambda_j \leq x\}; \qquad x \in \mathbb{R}, \tag{2.4}$$

counting the number of eigenvalues in the subset x included in $\mathbb{R}$, where # denotes the cardinality of the set. This ESD is the normalized counting measure of the eigenvalues. The study of convergence in the empirical spectral distributions (F_{A_n}) for a given sequence of random matrices (A_n) is a significant challenge in RMT. Bai and Silverstein (2010), and Yao et al. (2015b) stated that the importance of ESD is that many important statistics in multivariate analysis can be expressed as functionals of the ESD of some Random Matrices.

If F_A converges to a deterministic distribution function F as n tends to infinity, then $F(x)$ is called the limiting spectral distribution (LSD) of A. Marchenko and Pastur (1967) first derived the LSD of the sample covariance matrix and Bai and Silverstein (2004) studied the central limit theorem (CLT) for its linear spectral statistics (LSSs) defined as,

$$\frac{1}{n}\sum_{i=1}^{n} f(\lambda_i) = \int f(x)\, dF_A(x), \tag{2.5}$$

where $f(.)$ is a function on $\mathbb{R}^+$. Moreover, the LSD explains the first-order limits of these LSSs, while the CLT characterizes the second-order asymptotic. In classical probability theory, these two are analogous to the Law of large numbers and the Central limit theorem, as described by Bai and Silverstein (2010), and Li et al. (2023).

According to Bai and Silverstein (2010), beyond a dimension of 4, the eigenvalues of a matrix do not possess a closed form. Therefore, understanding them necessitates the use of unique procedures. Three key methods are used in this field, moment method, Stieltjes transform, and orthogonal polynomial decomposition, with the third method considering underlying distributions in the RM.

2.2.1 Moment method

The moment method in Random Matrix Theory (RMT) is a powerful technique used to analyze the statistical properties of matrices with random entries. It is an application of the Moment Convergence Theorem (MCT), which states conditions under which the moments of a sequence of random variables converge to those of a limiting distribution (Bai and Silverstein (2010), and Yao et al. (2015a)).

In the following, F_n will denote a sequence of distribution functions, and the k-th moment of the distribution F_n is denoted by

$$m_{n,k} = m_k(F_n) := \int x^k\, dF_n(x). \tag{2.6}$$

If we consider the ESD F_A of a matrix A ($n \times n$ Hermitian matrix). Then, the k-th moment of F_A is given by,

$$m_{n,k}(A) = \int x^k\, dF_A(x) = \frac{1}{n}\mathrm{tr}(A^k). \tag{2.7}$$

This relation (2.7) holds significant importance within Random Matrix Theory (RMT). As mentioned in Bai and Silverstein (2010), the problem of demonstrating the ESD of a sequence of random matrices $\{A_n\}$ converging to a limit (strongly, weakly, or in another sense) reduces to demonstrating that the sequence $\frac{1}{n}\mathrm{tr}(A^k)$ tends to a limit m_k in the corresponding sense for each fixed k, and then verifying the Carleman condition (2.8). Utilizing the Monotone Convergence Theorem enables the achievement of this objective.

2.2.1.1 Carleman condition

In analysis, Carleman's condition gives a sufficient condition for the determinacy of the moment problem. That is, no other measure ν has the same moments as μ if a measure μ meets Carleman's condition.

Let $\{m_k = m_k(F)\}$ be the sequence of moments of the distribution function F. If the Carleman condition

$$\sum_{k=1}^{\infty} {m_{2k}}^{-1/2k} = \infty, \tag{2.8}$$

is satisfied, then F is uniquely determined by the moment sequence $\{m_k, k = 0, 1, .\}$. In the majority of instances, the limiting spectral distribution (LSD) exhibits finite support. Many findings related to determining the LSD or establishing its existence have been derived through the estimation of the mean, variance, or higher moments of $\frac{1}{n}\mathrm{tr}(A^k)$ (Bai and Silverstein, 2010).

2.2.2 Stieltjes transform

Stieltjes transform is a mathematical tool used in the theory of functions of a complex variable and in various areas of applied mathematics, including probability theory and random matrix theory as presented in Bai and Silverstein (2010). Similar to how the characteristic function of a probability distribution becomes a useful tool for Central limit theorems, the Stieltjes transform is a practical and extremely effective technique in the study of the convergence of spectral distribution of matrices (or operators).

Definition 1. *(Bai and Silverstein, 2010) Let* $\mathbb{R}$ *be the real numbers and let* $\mathbb{C}$ *be the complex numbers. Let* $\mathcal{P}(\mathbb{R})$ *represent the set of probability measures on the real line, and let* $\mathcal{F}(\mathbb{C})$ *denote the set of functions defined on the complex plane. The Stieltjes transform, denoted by* $\mathcal{S}$ *is defined as a mapping from* $\mathcal{P}(\mathbb{R})$ *to* $\mathcal{F}(\mathbb{C})$*.*

$$S : \mathcal{P}(\mathbb{R}) \to \mathcal{F}(\mathbb{C}),$$

$$\mu \to m = S(\mu),$$

where $\mu \in \mathcal{P}(\mathbb{R})$ *is a probability measure and* m: $\mathbb{C} \backslash supp(\mu) \to \mathbb{C}$,

$$m(z) = \frac{1}{x-z}\mu(dx), \quad z \in \mathbb{C}^{+}. \tag{2.9}$$

Here $\mathbb{C}^{+}$ *denotes the complex numbers with positive imaginary part. The function* m *is called the Stieltjes function of* μ*.*

Let A be an $n \times n$ Hermitian matrix and F_n be its ESD. Then, the Stieltjes transform of F_n is given by,

$$s_n(z) = \int \frac{1}{x-z}\, dF_n(x) = \frac{1}{n}\mathrm{tr}\,(A - zI)^{-1}. \tag{2.10}$$

2.2.3 Orthogonal polynomial decomposition

Deift (2000) defined orthogonal polynomial decomposition as a technique used in Random Matrix Theory (RMT) to analyze the joint probability density function of the eigenvalues of random matrices. In the context of RMT, random matrices often represent complex systems with random interactions. Therefore, understanding the statistical properties of these matrices is crucial for gaining insights into the behavior of the systems they represent, as explained by Van Assche (2020).

Suppose that the matrix A follows a probability density $p_n = H(\lambda_1, ...\lambda_n)$. It is well established that the joint density function of its eigenvalues takes the form $p_n(\lambda_1, \cdots \lambda_n) = cJ(\lambda_1, \cdots \lambda_n)H(\lambda_1, \cdots \lambda_n)$. where J stems from the integration of the Jacobian of the transform from the matrix space to its eigenvalue-eigenvector space.

Typically, it is assumed that H adopts the structure $H(\lambda_1, \cdots \lambda_n) = \prod_{k=1}^{n} g(\lambda_k)$, and J takes the form $\prod_{i<j}(\lambda_i - \lambda_j)^\beta \prod_{k=1}^{n} h_n(\lambda_k)$. For example, $\beta = 1$ and $h_n = 1$ for a real Gaussian matrix, $\beta = 2$, $h_n = 1$ for a complex Gaussian matrix, and $\beta = 1$ and $h_n(x) = x^{n-p}$ for a real Wishart matrix with $n \geq p$. In this regard, Deift (2000), and Bai and Silverstein (2010) provide more detailed applications and examples.

2.2.4 Marčenko-Pastur law

The Marchenko-Pastur distribution (M-P), a cornerstone of random matrix theory, describes the asymptotic behavior of singular values in large rectangular random matrices. This well-known law was established in 1967 by mathematicians Volodymyr Marchenko and Leonid Pastur and is documented in Marchenko and Pastur (1967).

The M-P law $F_\gamma(x)$ has a density function

$$p_\gamma(x) = \begin{cases} \frac{1}{2\pi x\gamma\sigma^2}\sqrt{(b_\gamma - x)(x - a_\gamma)}, & \text{if } a_\gamma \leq x \leq b_\gamma, \\ 0, & \text{otherwise,} \end{cases}$$

and has a point mass $1 - 1/\gamma$ at the origin if $\gamma > 1$, where $a_\gamma = \sigma^2(1 - \sqrt{\gamma})^2$ and $b_\gamma = \sigma^2(1+\sqrt{\gamma})^2$. Here, the constant γ is the dimension to the sample size ratio index, and σ^2 is the scale parameter. If $\sigma^2 = 1$, the M-P law is said to be the standard M-P law as described by Bai and Silverstein (2010).

The moments of M-P law are (see Appendix (A.1.1)),

$$m_k(\gamma) = \int_{a_\gamma}^{b_\gamma} x^k d\sigma_{MP,\gamma}(x) = \sum_{r=0}^{k-1} \frac{1}{r+1}\binom{k}{r}\binom{k-1}{r}\gamma^r; \quad k \geq 1. \tag{2.11}$$

2.2.5 Path-Shortening algorithm PS(I)

Heiny and Mikosch (2018) defined the Path-shortening algorithm as a computational technique utilized in conjunction with the method of moments to derive accurate bounds for matrix norms, particularly in the context of sample correlation matrices. This approach effectively takes advantage of the inherent structure of these matrices to accelerate the computation, yielding accurate estimations of matrix norms.

The result of applying the path-shortening function PS to a path I is represented by $(S(I), runs(I), simples(I))$ as produced by the algorithm. Here, $S(I)$ signifies the resulting shortened path, $runs(I)$ denotes the overall count of vertices removed through **Type-I** reductions, and $simples(I)$ indicates the total number of vertices eliminated via **Type-II** reductions. Heiny and Mikosch (2018), and Heiny and Yao (2022) described the following as the output $PS(I)$ of the mentioned algorithm.

Input: Path $I = (i_1, \ldots, i_k)$. Set $J = I$ and $R = 0$,$runs = 0$.

Step 0: Set $l = \text{length}(I)$. Go to **Step 1**.

Step 1: Erase runs.

– If $i_j = i_{j+1}$ for some $1 \leq j \leq l$, where we interpret i_{l+1} as i_1, erase element i_j from the path. Set $I = (i_1, \ldots, i_{j-1}, i_{j+1}, \ldots, i_l)$, $runs = \text{runs} + 1$ and return to Step 0.

– Otherwise, proceed with **Step 2**.

Step 2: Let R_1 be the number of elements of the path I which appear exactly once. Set simples := simples + s. Then define I to be the resulting (possibly shorter) path obtained by deleting those s elements from the path I. Go to **Step 3**.

Step 3:

– If $J = I$, then return $(I, \text{runs}, \text{simples})$ as output.

– If $J \neq I$, set $J := I$ and return to **Step 0**.

For any path I, have the identity

$$|I| = |S(I)| + \text{runs}(I) + \text{simples}(I). \tag{2.12}$$

2.2.6 Some terminology from graph theory and combinatorics

To evaluate the limits of moments of the ESD of a sample correlation matrix, literature shows some useful information from combinatorics and graph theory. As given in Bai and Silverstein (2010), this is because the mean and variance of each empirical moment are expressed as the sum of the expectations of the products of matrix entries. To ensure accuracy, it is important to systematically account for the number of significant terms within these expressions. Thus, in this section, we introduce concepts from graph theory as explained by Bai and Silverstein (2010), and Heiny and Yao (2022).

The set of the first m positive integers $\{1, 2, \ldots, m\}$ is represented as $[[1, m]]$. A tuple $I = (i_1, i_2, \ldots, i_k) \in [[1, p]]^k$ of positive integers is defined as a path with vertices $i_l \in \{1, 2, \ldots, p\}$. The length of the path is denoted by $k = |I|$. The set of distinct elements in I is denoted by $\{I\}$. The cardinality of a set A is denoted by $\#A$. If $r = |\{I\}|$, then I is termed an r-path. For instance, if $I = (1, 1, 2, 2)$, it has a length of 4, and it is considered a 2-path since $\{I\} = \{1, 2\}$. A path is defined as canonical if $i_1 = 1$ and $i_l \leq \max\{i_1, i_2, \ldots, i_{l-1}\} + 1$ for $l \geq 2$. A canonical r-path I satisfies $\{I\} = [[1, r]]$ (Bai and Silverstein (2010), and Heiny and Yao (2022)).

To utilize the moment method for demonstrating the convergence of the ESD of high-dimensional sample correlation matrices to the M-P law, it is essential to grasp the characteristics of a specific category of Δ-graphs. Suppose that $i_1, \ldots, i_k$ are k positive integers (not necessarily distinct) not greater than p, and $j_1, \ldots, j_k$ are k positive integers (not necessarily distinct) not larger than n. A $\Delta-$ graph can be defined as follows. Draw two parallel lines, referred to as the I line and the J line. Plot $i_1, \ldots, i_k$ on the I line and $j_1, \ldots, j_k$ on the J line. Then, draw k (down) edges from i_u to j_u for $u = 1, \ldots, k$ and k (up) edges from j_u to i_{u+1} for $u = 1, \ldots, k$ (with the convention that $i_{k+1} = i_1$). The graph is denoted by $G(i, j)$, where $i = (i_1, \ldots, i_k)$ and $j = (j_1, \ldots, j_k)$. An example of such a graph in Heiny and Yao (2022) with $k = 3$, is given in Figure 2.1. Two graphs are considered isomorphic when they can be transformed into each other using appropriate permutations on the sets $(1, 2, \ldots, p)$ and $(1, 2, \ldots, n)$, and each isomorphism class has a unique canonical graph as provided in Bai and Silverstein (2010), Heiny and Mikosch (2018), and Heiny and Yao (2022).

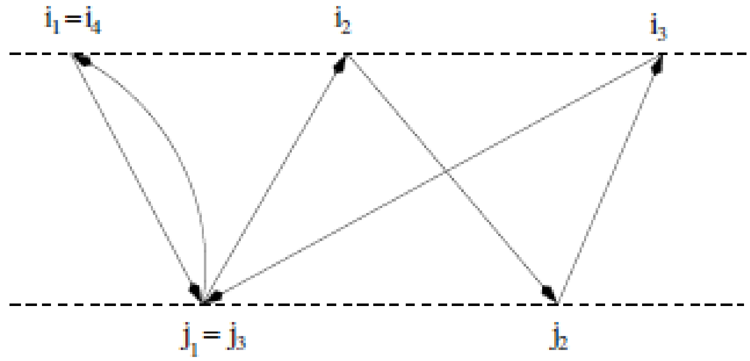

Figure 2.1: $\Delta-$ Graph with $I = (i_1, i_2, i_3)$ and $T = (j_1, j_2, j_1)$
.

2.3 Comparison and validation

The use of simulations to validate theoretical methods for estimating the limiting distributions of eigenvalues of sample correlation matrices is an essential stage in the process of comparison and confirmation. This ensures the precision and reliability of the produced limiting distributions, providing insights into the behavior of the sample correlation matrix's eigenvalues and their asymptotic properties (Heiny and Mikosch (2018), Heiny and Yao (2022), Jiang (2004), and El Karoui (2009)).

Chapter 3

Results and Discussion

This chapter discusses the various conditions required to derive the limiting distributions for the eigenvalue spectra of sample correlation matrices. The findings are categorized into four primary scenarios based on input data characteristics while providing insights into method behavior and performance across different conditions. The structure of this chapter is as follows, Section 3.1 focuses on the case of i.i.d entries with finite fourth moment. Section 3.2 addresses the case of i.i.d entries with infinite fourth moment and finite second moment. In Section 3.3, we discuss the case of i.i.d entries with infinite second moment. Finally, Section 3.4 explores the case of identically distributed, but dependent entries.

3.1 The case of i.i.d entries with finite fourth moment

The assumption of independent and identically distributed (i.i.d) entries with a finite fourth moment serves as a foundational premise, facilitating the utilization of specific statistical techniques designed for finite moments. The scenario implies that the elements in our dataset are i.i.d, with a bounded fourth moment. Heiny and Mikosch (2018) enhanced the findings of Jiang (2004) and proved that, application of Weyl's inequality by Bhatia (1997) for the sample correlation matrix $R = F^{1/2}XX'F^{1/2}$ with

$F = \mathrm{diag}(1/D_1, \dots, 1/D_p)$, where

$$D_i = D(n)_i = \sum_{j=1}^{n} X_{ij}^2, \quad i = 1, \dots, p; \quad n \geq 1.$$

yields,

$$\max_{i=1,\dots,p} \left|\lambda_i(R) - n^{-1}\lambda_i(S)\right| \leq \left|XX'F - n^{-1}XX'\right|_2 \leq n^{-1}\left|XX'\right|_2 |nF - I|_2 \qquad (3.1)$$
$$= n^{-1}\lambda_1(S) \max_{i=1,\dots,p} \left|\frac{nD_i}{n} - 1\right|.$$

For any matrix $\mathbf{A}$, $|\mathbf{A}|_2$ denotes its spectral norm, i.e., its largest singular value (see Appendix A3). Furthermore, leveraging the concept introduced in Lemma 2 of Bai and Yin (2008), it is established that the condition $\mathbb{E}[\xi^4] < \infty$ is equivalent to

$$\max_{i=1,\dots,p} |\frac{nD_i}{n} - 1| \xrightarrow[n\to\infty]{a.s.} 0.$$

Hence, as n tends to infinity, $\max_{i=1,\dots,p} |\lambda(R) - n^{-1}\lambda(S)| \to 0$ a.s.

Bao et al. (2012) assumed independently and identically distributed, symmetric entries X_{ij} and suggested the existence of positive constants C and C', where the $P(|X| \geq t^C) \leq e^{-t}, \quad t \geq C'$. Also, they showed that

$$n^{2/3} \frac{\left(\sqrt{\gamma}\right)^{1/3}}{\left(1+\sqrt{\gamma}\right)^{4/3}} \left(\frac{\lambda_1(R)}{n} - \left(1 + \frac{\sqrt{p}}{n}\right)^2\right) \xrightarrow[n\to\infty]{d} \xi. \qquad (3.2)$$

where the random variable in the limit follows a Tracy–Widom distribution of the first-order.

3.2 The case of i.i.d entries with infinite fourth moment and finite second moment

This section discusses the case where the entries are i.i.d but possess an infinite fourth moment. This circumstance requires particular approaches to effectively handle the heavy-tailed data.

When $\mathbb{E}[\xi^4] = \infty$, the approach used in relation (3.1) to sample correlation matrices becomes inadequate or inappropriate. Under the Theorem 2.3, Bai and Zhou (2008) proved that if the random variable is within the domain of attraction of the normal law, then as the $p/n \to c$ (approaches a constant value), F_R converges almost surely to the Marčenko-Pastur distribution. The Stieltjes transform used for the characterization is given by,

$$m_R(z) = -\left(cz - c + 1\right) + \frac{\left(cz - c - 1\right)^2 - 4c}{2z}, \quad z \in \mathbb{C}^+ \tag{3.3}$$

where $\mathbb{C}^+$ denotes the complex numbers with positive imaginary part.

Heiny and Mikosch (2018) demonstrated that the limiting spectral distribution of the sample correlation matrices converges to the Marčenko-Pastur law. Subsequently, Theorem 3.3 of the study shows that the extreme eigenvalues converge almost surely to the endpoints of the limiting support. The findings of Heiny and Mikosch (2018) are a refinement of the results of Jiang (2004). Both the aforementioned studies used the following basic assumptions,

$$E[Y_{11}Y_{12}] = o(n^{-2}) \quad \text{and} \quad E[Y_{11}^4] = o(n^{-1}), \quad n \to \infty, \tag{3.4}$$

and

$$\lim_{n\to\infty} \frac{p_n}{n} = \gamma \in (0, 1]. \tag{3.5}$$

Here $Y_{ij} = \frac{X_{ij}}{\sqrt{\sum_{j=1}^{n} X_{ij}^2}} \quad , \quad i = 1, \ldots, p \quad ; \quad j = 1, \ldots, n.$

Theorem 3.1 in Heiny and Mikosch (2018) states that if the centered random variable ξ satisfies the conditions in (3.4) and (3.5), then F_R converges weakly to the

Marchenko–Pastur law with index $\gamma > 0$. If ξ is symmetric and the condition from (3.4) does not hold, i.e., $\liminf_{n\to\infty} n\mathbb{E}[Y_{11}^4] > 0$, then $\liminf_{n\to\infty} \mathbb{E}\left[\int x^k F_R(dx)\right] > m_k(\gamma), \quad k \geq 1$. Here $m_k(\gamma)$ is the k-th moment of the M-P law given in (2.11). Hence, it is evident that the requirement

$$\liminf_{n\to\infty} n\mathbb{E}[Y_{11}^4] = 0, \tag{3.6}$$

is a necessary and sufficient condition for the convergence of F_R to the Marčenko-Pasture Law. Furthermore, Theorem 3.3 in Heiny and Mikosch (2018) mentions that under the above main assumption (3.5), (i) if $\mathbb{E}[\xi^4] < \infty$ and $\mathbb{E}[\xi] = 0$, (ii) or ξ is symmetric and satisfies the condition (A.2), then

$$\begin{aligned} \lambda_1(R) &\xrightarrow[n\to\infty]{a.s} (1+\sqrt{\gamma})^2 \\ \lambda_p(R) &\xrightarrow[n\to\infty]{a.s} (1-\sqrt{\gamma})^2. \end{aligned} \tag{3.7}$$

3.3 The case of i.i.d entries with infinite second moment

This section describes the case where the entries are i.i.d with the infinite second moment. Giné et al. (1997) suggested that the condition in (3.6) is satisfied when the distribution of ξ falls within the domain of attraction of normal law, which is equivalent to the function $E[\xi^2 \mathbf{1}_{\{|\xi|\leq x\}}]$ being slowly varying. Mason and Zinn (2005) further proved that the centered random variable ξ is regularly varying with the index $\alpha \in (0, 2)$ if and only if

$$\liminf_{n\to\infty} n\mathbb{E}[Y_{11}^4] = 1 - \frac{\alpha}{2}. \tag{3.8}$$

.

Heiny and Yao (2022) examine the breach of the condition (3.6), specifically indicating that $\mathbb{E}[\xi^2] = \infty$. Considering a symmetric distribution with a regularly varying tail behavior characterized by an index $\alpha \in (0, 2)$, Heiny and Yao (2022) demonstrated that the sequence of ESDs F_R converges weakly to a novel distribution denoted as $H_{\alpha,\gamma}$ termed the α-heavy Marčenko–Pastur law, with a parameter γ (Theorem 2.1). The class of distributions $H_{\alpha,\gamma}$ can be smoothly extended at the limits $\alpha = 2$ and $\alpha = 0$ resulting in the M-P law and a modified Poisson distribution (Theorem 2.2).

Ultimately, the condition in (3.6) is violated when ξ exhibits regular variation with the index $\alpha \in (0, 2)$. The findings presented in the aforementioned reference, reveals that under Theorem 2.1 each $H_{\alpha,\gamma}$ is uniquely characterized by the moment sequence $\mu_k(\alpha, \gamma) = \int x^k dH_\alpha, \gamma(x), k \leq 1$. The sequence $\mu_k(\alpha, \gamma)$ can be partitioned into two distinct components: the Marčenko-Pastur(M-P) part and a component associated with heavy tails, defined as follows,

$$\mu_k(\alpha, \gamma) = \begin{cases} m_k(\gamma), & \text{if } k = 1, 2, 3, \\ m_k(\gamma) + d_k(\alpha, \gamma), & \text{if } k \geq 4, \end{cases} \tag{3.9}$$

with $\quad d_4(\alpha, \gamma) = \left(1 - \frac{\alpha}{2}\right)^2 \gamma \quad$ and $\quad d_5(\alpha, \gamma) = \left(1 - \frac{\alpha}{2}\right)^2 (5\gamma + 5\gamma^2)$.

3.4 The case of identically distributed but dependent entries

The presence of dependency introduces additional complexity to the spectral analysis. Due to the complex structure of the sample correlation matrix, previous research has mostly examined situations where the sample data consists of independent components. This leads to a population covariance matrix that is diagonal and a correlation matrix that is an identity matrix, as given in Li et al. (2023). However, it is noteworthy that while these results provide valuable insights, their practical utility is constrained due to the excessively strict assumption of independence. This limitation significantly narrows down the applicability of such results in real-world settings. Moreover, there is a scarcity of research on the sample correlation matrix obtained from the dependent data (El Karoui (2009), Morales-Jimenez et al. (2021), and Li et al. (2023)). Aiming to address a gap in the literature, Li et al. (2023) researched the limiting spectral distribution of a high-dimensional Kendall's rank correlation matrix. The findings of the above reference revealed that the outcome deviates from the generalized Marčenko-Pastur law. Notably, the underlying population in their study encompasses a broad dependence structure, featuring high-dimensional correlated data. The data was treated as a sample $x_1, \ldots, x_n \in \mathbb{R}^p$ with the sign vector

$$A_{ij} = \operatorname{sign}(x_i - x_j) = (\operatorname{sign}(x_{i1} - x_{j1}), \ldots, \operatorname{sign}(x_{ip} - x_{jp}))^\top,$$

where sign(·) represents the sign function. Then, the sample Kendall's rank correlation matrix K_n is computed as:

$$K_n = \frac{2}{n(n-1)} \sum_{1 \le i < j \le n} A_{ij} A_{ij}^\top.$$

El Karoui (2009) suggests that the methods for sample covariance matrices can be extended to sample correlation matrices if $\mathbb{E}[\xi^4] < \infty$. This is supported by arguments that are comparable to relation (3.1). Furthermore, the findings in Theorem 1 of El Karoui (2009) reveal that matrices R and S converge asymptotically when the variable's fourth moment is finite and the spectral norm of R is uniformly bounded (i.e $|\mathbf{R}|_2 < K; K$ is a constant).

Chapter 4

Conclusion

This major paper provides a comprehensive overview of the limiting distributions for eigenvalues in sample correlation matrices from heavy-tailed populations, outlining key findings and their broader statistical implications.

As mentioned in Bai and Silverstein (2010), in practical scenarios where asymptotic theorems are applied to analyze the spectra of large-dimensional random matrices, two key issues arise following the identification of the LSD. The first issue revolves around bounding extreme eigenvalues, while the second focuses on understanding how quickly the Empirical Spectral Distribution (ESD) converges as the sample size increases (Bai and Silverstein, 2010). In general, LSD outlines the first-order limits of the LSSs while the CLT delineates their second-order asymptotic behaviors. Analogous to the law of large numbers and the central limit theorem in classical probability theory, the LSD and CLT play equivalent roles in characterizing the behavior of random matrices as presented by Li et al. (2023). The exploration of the limits of extreme eigenvalues holds significance not only for the applicability of the LSD in conjunction with the Helly-Bray theorem in Bai and Silverstein (2010) but also for its direct relevance in various practical domains. Bai and Silverstein (2010) points out that fields such as signal processing, pattern recognition, and edge detection rely on understanding the support of the LSD of population correlation matrices, which often comprise multiple disconnected regions.

Tracy and Widom (1996) discovered the expression for the maximum eigenvalue of a Gaussian matrix when appropriately normalized, and their work was the first to focus on the limiting distributions of extreme eigenvalues. Johnstone (2001) discovered the limiting distribution of the largest eigenvalue of a Wishart matrix, which plays a crucial role in the PCA. Jiang (2004) initially derived the LSD of the Pearson-type sample correlation matrix, while Heiny and Mikosch (2018) proposed a refinement of the same in 2018. Gao et al. (2017) further advanced the CLT for its LSSs. Bao et al. (2012) later established the Tracy-Widom law for its extreme eigenvalues, and Pillai and Yin (2012) subsequently extended this result to more general cases. Bai and Silverstein (2010) examined a particular instance of the CLT when the underlying distribution is complex Gaussian. In the recent literature, Heiny and Yao (2022) established that the sequence of ESDs F_R converges weakly to a new distribution $H_{\alpha,\gamma}$, termed the α-heavy M-P law with parameter γ. The study identified that the family of distributions $H_{\alpha,\gamma}$ has continuous extensions at the boundary, resulting in the standard M-P law and a modified Poisson distribution. However, all of these asymptotic results were derived under the assumption that the data samples were made up of independent components.

Recent focus has shifted towards non-parametric correlation matrices such as Kendall's τ and Spearman's ρ, particularly in handling heavy-tailed data samples characterized by a general dependent structure. These alternatives are rank-based, eliminating the necessity to impose any moment restrictions on the underlying distribution. Moreover, classical non-parametric statistics theory indicates that utilizing only the ranks of the data preserves robustness while sacrificing only partial information. Li et al. (2023) introduced a brand new finding on the above concept using Kendall's τ correlation matrix. A concise overview of the advancements related to the sample correlation matrix and Kendall's τ are presented in Table A.1.

Bibliography

Adhikari, S. and Friswell, M. (2007). Random matrix eigenvalue problems in structural dynamics. *International Journal for Numerical Methods in Engineering*, 69(3):562–591.

Ahmad, Z., Mahmoudi, E., Hamedani, G., and Kharazmi, O. (2020). New methods to define heavy-tailed distributions with applications to insurance data. *Journal of Taibah University for Science*, 14(1):359–382.

Anderson, T. W. (1963). Asymptotic Theory for Principal Component Analysis. *The Annals of Mathematical Statistics*, 34(1):122 – 148.

Arfken, G. (1985). *Mathematical Methods for Physicists*. Academic Press, 3rd edition.

Arnold, L. (1967). *On the asymptotic distribution of the eigenvalues of random matrices*. Mathematics Research Center, United States Army, the University of Wisconsin.

Arnold, L. (1971). On wigner's semicircle law for the eigenvalues of random matrices. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, 19(3):191–198.

Arous, G. B. and Guionnet, A. (2008). The spectrum of heavy tailed random matrices. *Communications in Mathematical Physics*, 278(3):715–751.

Auffinger, A., Ben Arous, G., and Péché, S. (2009). Poisson convergence for the largest eigenvalues of heavy tailed random matrices. In *Annales de l'IHP Probabilités et statistiques*, volume 45, pages 589–610.

Bai, Z. and Silverstein, J. W. (2010). *Spectral analysis of large dimensional random matrices*, volume 20. Springer.

Bai, Z. and Zhou, W. (2008). Large sample covariance matrices without independence structures in columns. *Statistica Sinica*, pages 425–442.

Bai, Z. D. and Silverstein, J. W. (2004). CLT for linear spectral statistics of large-dimensional sample covariance matrices. *The Annals of Probability*, 32(1A):553 – 605.

Bai, Z. D. and Silverstein, J. W. (2008). Clt for linear spectral statistics of large-dimensional sample covariance matrices. In *Advances In Statistics*, pages 281–333. World Scientific.

Bai, Z. D. and Yin, Y. Q. (1988a). Convergence to the semicircle law. *The Annals of Probability*, 16(2):863–875.

Bai, Z. D. and Yin, Y. Q. (1988b). Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a wigner matrix. *The Annals of Probability*, 16(4):1729–1741.

Bai, Z.-D. and Yin, Y.-Q. (2008). Limit of the smallest eigenvalue of a large dimensional sample covariance matrix. In *Advances In Statistics*, pages 108–127. World Scientific.

Baik, J. and Silverstein, J. W. (2006). Eigenvalues of large sample covariance matrices of spiked population models. *Journal of multivariate analysis*, 97(6):1382–1408.

Bandeira, A. S., Lodhia, A., and Rigollet, P. (2017). Marčenko-pastur law for kendall's tau.

Bao, Z. (2019). Tracy–widom limit for kendall's tau. *The Annals of Statistics*, 47(6):3504–3532.

Bao, Z., Pan, G., and Zhou, W. (2012). Tracy-widom law for the extreme eigenvalues of sample correlation matrices.

Basrak, B., Cho, Y., Heiny, J., and Jung, P. (2021). Extreme eigenvalue statistics of m-dependent heavy-tailed matrices. In *Annales de l'Institut Henri Poincare (B) Probabilites et statistiques*, volume 57, pages 2100–2127. Institut Henri Poincaré.

Belinschi, S., Dembo, A., and Guionnet, A. (2009). Spectral measure of heavy tailed band and covariance random matrices. *Communications in Mathematical Physics*, 289(3):1023–1055.

Bhati, D. and Ravi, S. (2018). On generalized log-moyal distribution: a new heavy tailed size distribution. *Insurance: Mathematics and Economics*, 79:247–259.

Bhatia, R. (1997). *Matrix Analysis*, volume 169 of *Graduate Texts in Mathematics*. Springer-Verlag, New York.

Bilodeau, M. and Duchesne, P. (2002). Principal component analysis from the multivariate familial correlation matrix. *Journal of Multivariate Analysis*, 82(2):457–470.

Boik, R. J. (2003). Principal component models for correlation matrices. *Biometrika*, 90(3):679–701.

Brox, D. S. (2023). Singular spectrum analysis notes.

Cai, T. T. and Ma, Z. (2013). Optimal hypothesis testing for high dimensional covariance matrices. *Bernoulli*, 19(5B):2359 – 2388.

Couillet, R. and Debbah, M. (2011). *Random matrix methods for wireless communications*. Cambridge University Press.

Davis, A. (1977). Asymptotic theory for principal component analysis: Non-normal case 1. *Australian Journal of Statistics*, 19(3):206–212.

Davis, R. A., Heiny, J., Mikosch, T., and Xie, X. (2016). Extreme value analysis for the sample autocovariance matrices of heavy-tailed multivariate time series. *Extremes*, 19:517–547.

Deift, P. (2000). *Orthogonal Polynomials and Random Matrices: A Riemann-Hilbert Approach: A Riemann-Hilbert Approach*, volume 3. American Mathematical Soc.

El Karoui, N. (2007). Tracy–widom limit for the largest eigenvalue of a large class of complex sample covariance matrices.

El Karoui, N. (2009). Concentration of measure and spectra of random matrices: Applications to correlation matrices, elliptical distributions and beyond.

Gao, J., Han, X., Pan, G., and Yang, Y. (2017). High dimensional correlation matrices: The central limit theorem and its applications. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 79(3):677–693.

Giné, E., Götze, F., and Mason, D. M. (1997). When is the Student t-statistic asymptotically standard normal? *The Annals of Probability*, 25(3):1514 – 1531.

Guillen, M., Prieto, F., and Sarabia, J. M. (2011). Modelling losses and locating the tail with the pareto positive stable distribution. *Insurance: Mathematics and Economics*, 49(3):454–461.

Heiny, J. and Mikosch, T. (2018). Almost sure convergence of the largest and smallest eigenvalues of high-dimensional sample correlation matrices. *Stochastic Processes and their Applications*, 128(8):2779–2815.

Heiny, J. and Yao, J. (2022). Limiting distributions for eigenvalues of sample correlation matrices from heavy-tailed populations. *The Annals of Statistics*, 50(6):3249–3280.

Hoyle, D. C. and Rattray, M. (2004). Principal-component-analysis eigenvalue spectra from data with symmetry-breaking structure. *Physical Review E*, 69(2):026124.

Jiang, T. (2004). The limiting distributions of eigenvalues of sample correlation matrices. *Sankhyā: The Indian Journal of Statistics*, pages 35–48.

Johnstone, I. M. (2001). On the distribution of the largest eigenvalue in principal components analysis. *The Annals of statistics*, 29(2):295–327.

Johnstone, I. M. and Yang, J. (2018). Notes on asymptotics of sample eigenstructure for spiked covariance models with non-gaussian data. *arXiv preprint arXiv:1810.10427*.

Jolliffe, I. T. and Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences*, 374(2065):20150202.

Kollo, T. and Neudecker, H. (1993). Asymptotics of eigenvalues and unit-length eigenvectors of sample variance and correlation matrices. *Journal of Multivariate Analysis*, 47(2):283–300.

Konishi, S. (1979). Asymptotic expansions for the distributions of statistics based on the sample correlation matrix in principal component analysis. *Hiroshima Mathematical Journal*, 9(3):647–700.

Ledoit, O. and Wolf, M. (2002). Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size. *The Annals of Statistics*, 30(4):1081 – 1102.

LePage, R. (2011). *Heavy-Tailed Distributions*, pages 626–628. Springer Berlin Heidelberg, Berlin, Heidelberg.

Li, W. and Qin, Y. (2014). Hypothesis testing for high-dimensional covariance matrices. *Journal of Multivariate Analysis*, 128:108–119.

Li, Z., Wang, C., and Wang, Q. (2023). On eigenvalues of a high-dimensional kendall's rank correlation matrix with dependence. *Science China Mathematics*, 66(11):2615–2640.

Lindsey, J. (2004). Cambridge series in statistical and probabilistic mathematics.

Marchenko, V. A. and Pastur, L. A. (1967). Distribution of eigenvalues for some sets of random matrices. *Matematicheskii Sbornik*, 114(4):507–536.

Mason, D. and Zinn, J. (2005). When does a randomly weighted self-normalized sum converge in distribution?

Morales-Jimenez, D., Johnstone, I. M., McKay, M. R., and Yang, J. (2021). Asymptotics of eigenstructure of sample correlation matrices for high-dimensional spiked models. *Statistica Sinica*, 31(2):571.

Parvin, C. A. (2004). An introduction to multivariate statistical analysis, 3rd ed. t.w. anderson. hoboken, nj: John wiley & sons, 2003, 742 pp., $99.95, hardcover. isbn 0-471-36091-0. *Clinical Chemistry*, 50:981–982.

Pastur, L. A. (1972). On the spectrum of random matrices. *Teoreticheskaya i Matematicheskaya Fizika*, 10(1):102–112.

Paul, D. (2007). Asymptotics of sample eigenstructure for a large dimensional spiked covariance model. *Statistica Sinica*, pages 1617–1642.

Pillai, N. S. and Yin, J. (2012). Edge universality of correlation matrices. *The Annals of Statistics*, pages 1737–1763.

Pillai, N. S. and Yin, J. (2014). Universality of covariance matrices. *The Annals of Applied Probability*, 24(3):935 – 1001.

Resnick, S. I. (1997). Heavy tail modeling and teletraffic data. *The Annals of Statistics*, pages 1805–1849.

Tikhomirov, K. (2015). The limit of the smallest singular value of random matrices with i.i.d. entries. *Advances in Mathematics*, 284:1–20.

Tracy, C. A. and Widom, H. (1996). On orthogonal and symplectic matrix ensembles. *Communications in Mathematical Physics*, 177:727–754.

Van Assche, W. (2020). Orthogonal and multiple orthogonal polynomials, random matrices, and painlevé equations. In *Orthogonal Polynomials: 2nd AIMS-Volkswagen Stiftung Workshop, Douala, Cameroon, 5-12 October, 2018 2*, pages 629–683. Springer.

Waternaux, C. M. (1976). Asymptotic distribution of the sample roots for a nonnormal population. *Biometrika*, 63(3):639–645.

Wigner, E. P. (1958). On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, 67(2):325–327.

Xiao, H. and Zhou, W. (2010). Almost sure limit of the smallest eigenvalue of some sample correlation matrices. *Journal of Theoretical Probability*, 23(1):1–20.

Yang, X., Liu, W., Liu, W., and Tao, D. (2019). A survey on canonical correlation analysis. *IEEE Transactions on Knowledge and Data Engineering*, 33(6):2349–2368.

Yao, J., Zheng, S., and Bai, Z. (2015a). *Large Sample Covariance Matrices and High-Dimensional Data Analysis*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press.

Yao, J., Zheng, S., and Bai, Z. (2015b). Sample covariance matrices and high-dimensional data analysis. *Cambridge UP, New York*.

Appendix A

Some Important Concepts and Additional Details

A.1 Marčenko-Pastur law

The Three standard M-P density functions for $\gamma \in \left\{\frac{1}{8}, \frac{1}{4}, \frac{1}{2}\right\}$ are displayed on Figure A.1. In particular, the density function behaves as $\sqrt{x - a_\gamma}$ and $\sqrt{b_\gamma - x}$ at the boundaries a_γ and b_γ, respectively (see Bai and Silverstein (2010)).

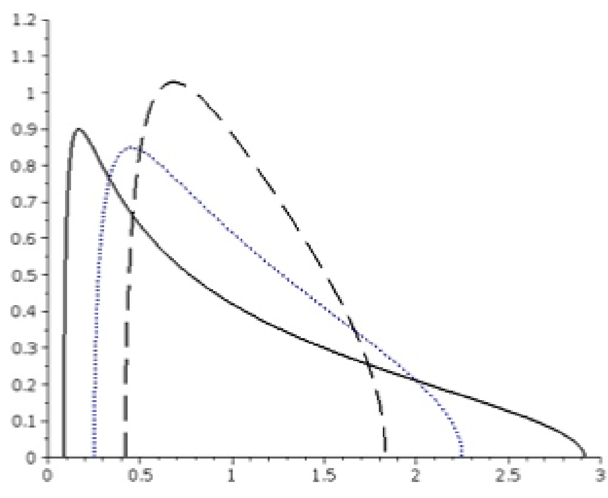

Figure A.1: Density plots of the Marčenko-Pastur distributions with indexes $\gamma = 1/8$ (dashed line), 1/4 (dotted line) and 1/2 (solid line).

.

A.1.1 Moments of M-P law

Proof. By definition,

$$m_k = \frac{1}{2\pi\gamma}\int_a^b x^{k-1}\sqrt{(b-x)(x-a)}\,dx, \quad \text{with } a=(1-\sqrt{\gamma})^2 \quad \text{and} \quad b=(1+\sqrt{\gamma})^2.$$

Let $x = 1+\gamma+z$. Then

$(x-a) = (x-(1-\sqrt{\gamma})^2) = (1+\gamma+z-(1-\sqrt{\gamma})^2) = (2\sqrt{\gamma}+z)$

$(b-x) = ((1+\sqrt{\gamma})^2-x) = ((1+\sqrt{\gamma})^2-1-\gamma-z) = (2\sqrt{\gamma}-z).$

Therefore, $(b-x)(x-a) = (2\sqrt{\gamma}-z)(2\sqrt{\gamma}+z) = (4\gamma-z^2).$

The corresponding boundaries are,

when $x=a \to x-a = 2\sqrt{\gamma}+z \to z=-2\sqrt{\gamma}$ and

when $x=b \to b-x = 2\sqrt{\gamma}-z \to z=2\sqrt{\gamma}.$

Then, we get

$$\begin{aligned} m_k &= \frac{1}{2\pi\gamma}\int_{-2\sqrt{\gamma}}^{2\sqrt{\gamma}}(1+\gamma+z)^{k-1}\sqrt{4\gamma-z^2}\,dz \\ &= \frac{1}{2\pi\gamma}\sum_{\ell=0}^{k-1}\binom{k-1}{\ell}(1+\gamma)^{k-1-\ell}\int_{-2\sqrt{\gamma}}^{2\sqrt{\gamma}} z^\ell\sqrt{4\gamma-z^2}\,dz. \end{aligned}$$

Now, let $z=2\sqrt{\gamma}u,$ then $dz = 2\sqrt{\gamma}du.$

The corresponding boundaries are, $u=-1$ and $u=1.$

Therefore, we get

$$\begin{aligned} m_k &= \frac{1}{2\pi\gamma}\sum_{\ell=0}^{k-1}\binom{k-1}{\ell}(1+\gamma)^{k-1-\ell}\int_{-1}^{1}(2u\gamma^{1/2})^\ell\sqrt{4\gamma-4\gamma u^2}\times 2\sqrt{\gamma}\,du \\ &= \sum_{\ell=0}^{k-1}\binom{k-1}{\ell}(1+\gamma)^{k-1-\ell}2^\ell\gamma^{\ell/2}\frac{2}{\pi}\int_{-1}^{1}u^\ell\sqrt{1-u^2}\,du. \end{aligned}$$

Now, let β_k denote the $k-$th moment of the Wigner's semicircular Law, because of the symmetry, the moments of the semicircle law are given by

$$\beta_k = \begin{cases} 0, & \text{if } k \text{ is odd,} \\ \mathbf{C}_{k/2}, & \text{if } k \text{ is even,} \end{cases}$$

where $\mathbf{C}_k$ is the $k-$th Catalan number, $\mathbf{C}_k = \frac{1}{k+1}\left(\frac{1}{2}\right)^{2k}\binom{2k}{k}.$

Therefore, we have

$$
\begin{aligned}
m_k &= \sum_{\ell=0}^{\lfloor(\frac{k-1}{2})\rfloor} \binom{k-1}{2\ell}(1+\gamma)^{k-1-2\ell}2^{2\ell}\gamma^{\ell} \quad \times \frac{1}{\ell+1}\left(\frac{1}{2}\right)^{2\ell}\binom{2\ell}{\ell} \\
&= \sum_{\ell=0}^{\lfloor(\frac{k-1}{2})\rfloor} \frac{(k-1)!}{\ell!(\ell+1)!(k-1-2\ell)!}\gamma^{\ell}(1+\gamma)^{k-1-2\ell} \\
&= \sum_{\ell=0}^{\lfloor(\frac{k-1}{2})\rfloor} \frac{(k-1)!}{\ell!(\ell+1)!(k-1-2\ell)!}\gamma^{\ell}\sum_{s=0}^{k-1-2\ell}\binom{k-1-2\ell}{s}\gamma^{s} \\
&= \sum_{\ell=0}^{\lfloor(\frac{k-1}{2})\rfloor}\sum_{s=0}^{k-1-2\ell} \frac{(k-1)!}{\ell!(\ell+1)!s!(k-1-2\ell-s)!}\gamma^{\ell+s}.
\end{aligned}
$$

Let $\ell + s = r$. Then, we get

$$
\begin{aligned}
m_k &= \sum_{\ell=0}^{\lfloor(\frac{k-1}{2})\rfloor}\sum_{r=\ell}^{k-1-\ell} \frac{(k-1)!}{\ell!(\ell+1)!(r-\ell)!(k-1-r-\ell)!}\gamma^{r} \\
&= \sum_{\ell=0}^{\lfloor(\frac{k-1}{2})\rfloor}\sum_{r=\ell}^{k-1-\ell} \frac{k(k-1)!(k-r)!r!}{k(k-r)!(\ell+1)!r!\ell!(r-\ell)!(k-1-r-\ell)!}\gamma^{r} \\
&= \frac{1}{k}\sum_{\ell=0}^{\lfloor(\frac{k-1}{2})\rfloor}\sum_{r=\ell}^{k-1-\ell} \frac{k!}{(k-r)!r!}\times\frac{(k-r)!}{(\ell+1)!(k-1-r-\ell)!}\times\frac{r!}{(r-\ell)!\ell!}\gamma^{r} \\
&= \frac{1}{k}\sum_{\ell=0}^{\lfloor(\frac{k-1}{2})\rfloor}\sum_{r=\ell}^{k-1-\ell}\binom{k}{r}\binom{k-r}{k-1-r-\ell}\binom{r}{\ell}\gamma^{r}.
\end{aligned}
$$

We can find the upper bound of ℓ as follows,

$\ell \le r \le k-1-\ell \quad \rightarrow \quad \ell \le k-1-r$ and $\ell \le r$.

Therefore, we have $\ell \le \min(r, k-1-r)$. Then, we get

$$
m_k = \frac{1}{k}\sum_{r=0}^{k-1}\binom{k}{r}\gamma^{r}\sum_{\ell=0}^{\min(r,k-1-r)}\binom{k-r}{k-1-r-\ell}\binom{r}{\ell}.
$$

When $\min(r, k-1-r) = k-1-r$, using Vandermonde's identity, we get

$$
\sum_{\ell=0}^{k-1-r}\binom{k-r}{k-1-r-\ell}\binom{r}{\ell} = \binom{k}{k-1-r} = \binom{k}{r+1}.
$$

When $\min(r, k-1-r) = r$, using Vandermonde's identity, we get

$$\begin{aligned}\sum_{\ell=0}^{r}\binom{k-r}{k-1-r-\ell}\binom{r}{\ell} &= \sum_{\ell=0}^{r}\binom{k-r}{\ell+1}\binom{r}{r-\ell}\\ &= \sum_{\ell=0}^{r}\frac{k-r}{\ell+1}\binom{k-r-1}{\ell}\binom{r}{r-\ell}\\ &= (k-r)\sum_{\ell=0}^{r}\frac{1}{\ell+1}\binom{r}{r-\ell}\binom{k-r-1}{\ell}\\ &= (k-r)\sum_{\ell=0}^{r}\frac{1}{r+1}\binom{r+1}{r-\ell}\binom{k-r-1}{\ell}\\ &= \frac{k-r}{r+1}\times\binom{k}{r} = \binom{k}{r+1}.\end{aligned}$$

Therefore, we have

$$m_k = \frac{1}{k}\sum_{r=0}^{k-1}\binom{k}{r}\binom{k}{r+1}\gamma^r = \sum_{r=0}^{k-1}\frac{1}{r+1}\binom{k}{r}\binom{k-1}{r}\gamma^r$$

as required.

A.2 Condition (G_q)

This condition is crucial for the proofs in Heiny and Mikosch (2018).

There exists a sequence $q = q_n \to \infty$ such that for some integer sequence $k = k_n$ with $k/\log n \to \infty$ we have $(k^3 q)/n \to 0$, and the moment inequality

$$\mathbb{E}\left[Y_1^{2m_1} \cdots Y_r^{2m_r}\right] \leq \frac{q_n}{n}\mathbb{E}\left[Y_1^{2m_1} \cdots Y_{r-1}^{2m_{(r-1)}} Y_r^{(2m_r-2)}\right]$$

holds for $1 \leq r \leq \ell - 1$ and any positive integers $m_1, \ldots, m_r$ satisfying $m_1 + \cdots + m_r = \ell$, where $\ell \leq k$.

A.3 Some concepts about matrices

A.3.1 Hermitian matrix

Definition 2. *(Arfken (1985)) A square matrix is called Hermitian if it is self-adjoint. Therefore, a Hermitian matrix $A = (a_{ij})$ is defined as one for which*

$$A = A^\dagger,$$

where $A^\dagger$ denotes the conjugate transpose. This is equivalent to the condition

$$a_{ij} = \overline{a_{ji}},$$

where $\overline{z} = (a - bi)$ denotes the complex conjugate of $z = (a + bi)$. As a result of this definition, the diagonal elements a_{ii} of a Hermitian matrix are real numbers (since $a_{ii} = \overline{a_{ii}}$), while other elements may be complex. Therefore, Hermitian matrices can be understood as the complex extension of real symmetric matrices.

A.3.2 Singular values of a matrix

Definition 3. *(Brox (2023)) Let A be an $m \times n$ matrix and consider the matrix $A^T A$. This is a symmetric $n \times n$ matrix, so its eigenvalues are real. Then the numbers*

*$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$ defined below are called the **singular values** of A.*

Lemma A.3.1. *(Brox (2023)) If λ is an eigenvalue of $A^T A$, then $\lambda \geq 0$.*

Proof. Let x be an eigenvector of $A^T A$ with eigenvalue λ. We compute that

$$\|Ax\|^2 = (Ax) \cdot (Ax) = (Ax)^T Ax = x^T A^T Ax = x^T(\lambda x) = \lambda x^T x = \lambda \|x\|^2.$$

Since $\|Ax\|^2 \geq 0$, it follows from the above equation that $\lambda\|x\|^2 \geq 0$. Since $\|x\|^2 > 0$ (since the convention is that eigenvectors are nonzero), we deduce that $\lambda \geq 0$.

Let $\lambda_1, \ldots, \lambda_n$ denote the eigenvalues of $A^T A$, with repetitions. Order these so that $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$. Let $\sigma_i = \sqrt{\lambda_i}$, so that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$. □

Proposition A.3.2. *(Brox (2023)) Let A be an $m \times n$ matrix. Then the maximum value of $\|Ax\|$, where x ranges over unit vectors in $\mathbb{R}^n$, is the largest singular value σ_1, and this is achieved when x is an eigenvector of $A^T A$ with eigenvalue σ_1^2.*

Proof. Let $v_1, \ldots, v_n$ be an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors of $A^T A$ with eigenvalues σ_i^2. If $x \in \mathbb{R}^n$, then we can expand x in this basis as

$$x = c_1 v_1 + \cdots + c_n v_n \tag{A.1}$$

for scalars $c_1, \ldots, c_n$. Since x is a unit vector, $\|x\|^2 = 1$, which (since the vectors $v_1, \ldots, v_n$ are orthonormal) means that

$$c_1^2 + \cdots + c_n^2 = 1.$$

On the other hand,

$$\|Ax\|^2 = (Ax) \cdot (Ax) = (Ax)^T(Ax) = x^T A^T Ax = x \cdot \left(A^T Ax\right).$$

By (A.1), since v_i is an orthonormal basis consisting eigen vectors of $A^T A$ with eigenvalue σ_i^2, we have

$$A^T A x = c_1 \sigma_1^2 v_1 + \cdots + c_n \sigma_n^2 v_n.$$

Taking the dot prodoct with (A.1), and using the fact that the vectors $v_1, \ldots, v_n$ are orthonormal, we get

$$\|Ax\|^2 = x \cdot \left(A^T A x\right) = \sigma_1^2 c_1^2 + \cdots + \sigma_n^2 c_n^2.$$

Since σ_1 is the largest singular value, we get

$$\|Ax\|^2 \leq \sigma_1^2 \left(c_1^2 + \cdots + c_n^2\right).$$

Equality holds when $c_1 = 1$ and $c_2 = \cdots = c_n = 0$. Thus, the maximum value of $\|Ax\|^2$ for a unit vector x is σ_1^2, which is achieved when $x = v_1$. □

A.4 Development of sample correlation matrices in the RMT.

Table A.1: Developments of sample correlation matrices in the random matrix theory

	Sample correlation	Kendall's τ correlation
Independent case ($\Sigma = I$)		
LSD	Jiang (2004); Heiny and Yao (2022); Heiny and Mikosch (2018); Morales-Jimenez et al. (2021)	Bandeira et al. (2017)
CLT for LSSs	Gao et al. (2017); Baik and Silverstein (2006)	Gao et al. (2017)
Tracy-Widom	Bao et al. (2012)	Bai and Silverstein (2008)
Dependent case (general Σ)		
LSD	El Karoui (2009)	Li et al. (2023)
CLT for LSSs	Heiny and Yao (2022)	
Tracy-Widom		

Vita Auctoris

NAME:	Sachini Wijesundara
PLACE OF BIRTH:	Sri Lanka
YEAR OF BIRTH:	1992
EDUCATION:	University of Peradeniya, B.Sc. in Mathematics and Statistics, Sri Lanka, 2014-2017.
	University of Peradeniya, M.Sc. in Applied Statistics, Sri Lanka, 2019-2021.
	University of Windsor, M.Sc in Statistics, Windsor, Ontario, 2022-2024.