Zhao_Zihui_01051725

by Zihui Zhao

Submission date: 06-Sep-2022 05:23AM (UTC+0100)
Submission ID: 185732073
File name: Zhao_Zihui_01051725.pdf (525.5K)
Word count: 11338
Character count: 48574

Imperial College London

Imperial College London

Department of Mathematics

Random Matrix Theory: Moment expansions, algebra and combinatorics

Author: Zihui Zhao (CID: 01051725)

A thesis submitted for the degree of

MSc in Mathematics and Finance, 2021-2022

Declaration

The work contained in this thesis is my own work unless otherwise stated.

Acknowledgements

I would like to express my sincere gratitude to my supervisor Professor Richard Martin, who has given me this great opportunity to learn about this fascinating subject. I appreciate your patience, valuable advice, and feedback throughout this project.

It has been a pleasure studying with my peers and friends, and I would like to thank them for their support.

Finally, I wish to express my gratitude to my mother for her unwavering support.

In memory of my father

Abstract

The main aim of this thesis is to find the relationship between the moments of the empirical covariance matrix E and the moments of the true covariance matrix C. In chapter 2, we treat some free probability theory and then review the results given by Bouchaud & Potters. In chapter 3, we provide the details to the diagrammatic method given by Sengupta & Mitra's paper. Then, we give a recurrence formula to compute the moments of the empirical covariance matrix E based on their results. In chapter 4, we developed the diagrammatic method further and give explicit formulae to compute the moments based on results from combinatorics and representation theory.

Contents

Chapter 1

Introduction

In many situations, we want to know the correlations between multiple variables. It is typically done by computing the empirical covariance matrix $E = \frac{XX^T}{T}$ from the data matrix X, where X is $N \times T$, N is the number of the variables and T is the number of observations.[1] In the above arguments, we actually assume that there is a true covariance matrix C, such that the data in matrix X satisfies

$$\mathbb{E}(X_{it}X_{js}) = C_{ij}. \tag{1.1}$$

Under this assumption, we know if we have an infinite number of observations for a fixed number of variables, that is the ratio $q = N/T$ tends to zero, then E converges almost surely to the true covariance matrix C. In real world, we do not have infinite data, but we are able to use a huge amount of data to achieve a very good pointwise estimation of C through equation (1.1). Sometimes, the number of variables that we are interested in are also very large such that $0 < q = N/T < 1$ is not very close to 0.

In this case, equation (1.1) still have a good estimation, but if we think of E as a whole, we shall observe some bias effects. These bias effects could be presented by comparing the eigenvalues of the empirical covariance matrix E and the true covariance matrix C. Since we are dealing with covariance matrices, so they are symmetric positive semi-definite. Therefore, the eigenvalues of covariance matrices are distributed on the positive part of the real-line. Suppose $\lambda_1, \ldots, \lambda_N$ are the eigenvalues, then we could define

$$F(x) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}_{\{x \leq \lambda_i\}}.$$

If we think F as a cumulative distribution function, then it defines a probability measure on $\mathbb{R}$, say μ. We call such probability μ as the spectrum of the covariance matrix. By computing the spectrum of E and C, we could see the bias effect. In figure 1.1, we illustrate several cases.

Also in the figure 1.1, we observed that when N gets large, the spectrum seems to look continuous. This motivates us to study the behaviour of the spectrum when N tends to infinity, as it would be much easier to study the continuous spectrum. So for a covariance matrix with large N, we could use the continuous spectrum to describe the spectrum of a covariance matrix.

In order to study the behaviour when N tends to infinity, we need to use the random matrix theory. Roughly speaking, random matrix is a matrix such that all its entries are random variables, or equivalently, a matrix-valued random variable. The empirical covariance matrix E, is a typical random matrix. We could thought all the entries of the data matrix X as a random variable, and hence $E = \frac{XX^T}{T}$ is a $N \times N$ random matrices. The true covariance matrix C could be thought as a constant random matrix.

An important tool from random matrix theory is called the Stieltjes transform. In brief, for a

[1] We assume that the data matrix X has already been centered.

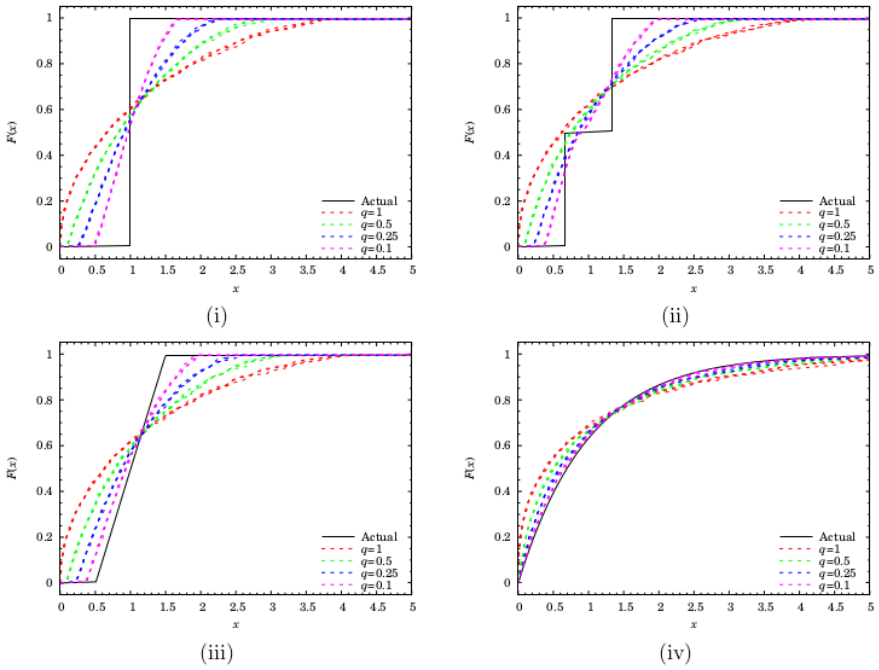


Figure 1.1: Empirical cdf (cumulative distribution function) compared with actual, for different models and different values of $q = N/T$ as indicated. In each case three realisations of the random matrix are shown, indicating some dispersion around a theoretical mean, as follows. (i) equal; (ii) $\frac{2}{3}$ or $\frac{4}{3}$ with equal probability; (iii) uniformly distributed on $[\frac{1}{2}, \frac{3}{2}]$; (iv) exponentially distributed of unit mean. All the models exhibit the same basic pattern, in which the highest eigenvalues are overestimated and the lowest ones underestimated (i.e. empirical spectrum has too much dispersion around the actual spectrum), with the effect decaying rather slowly as q is reduced to zero.

random matrix A, the Stieltjes transform of A equals to

$$\sum_{k=0}^{\infty} \frac{\tau(A^k)}{z^{k+1}},$$

where $\tau(A^k)$ is called the k-th moment of A. The k-th moment is given by

$$\tau(A^k) = \frac{1}{N}\mathrm{tr}\,(A^k) = \sum_{i=1}^{N} \lambda_i^k,$$

where λs are the eigenvalues of A. Note that in this case, λs are all random variables. The Stieltjes transform, if it is converged, can be used to compute the spectrum of the random matrix, which will be discussed in chapter 2. If we are interested in the relationship between the spectrum of E and the spectrum of C, since Stieltjes transform encodes the information about the spectrum, we could study the relationship between the moments of E and the moments of C. This is the main topic of this article.

As we have mentioned, the main object that we are interested in are the empirical covariance matrices E. More specifically, we assume that the data matrix satisfies:

$$\mathbb{E}(X_{it}X_{js}) = C_{ij}D_{ts}. \tag{1.2}$$

This is a more general form compare to the equation (1.1), we can get equation (1.1) by setting D to be the identity matrix. The matrix D is called the true temporal covariance matrix, in this

case, C is also called the true spatial covariance matrix. Another important assumption that we are making is all the entries X_it of X is Gaussian distributed. This property allows us to use the Wick's theorem to study the moment, and will lead us to some interesting results.

In chapter 2, we will first treat some free probability theories. Free probability theory is the generalization of the classical probability theory, which studies non-commutative random variables, for example, random matrices. We will define and study all the terminologies and tools we mentioned above in the language of free probability. Another important reason to use the free probability theory is the concept of freeness. Freenees is the generalization of the concept of independence from the classical probability theory. If the data matrix X satisfies the equation (1.2), then we can write the empirical covariance matrix E by

$$E = \frac{C^{\frac{1}{2}} X_0 D X_0^T C^{\frac{1}{2}}}{T}, \tag{1.3}$$

where the entries of X_0 are independent Gaussian random variables. To decompose this structure, we need to use the property of free non-commutative random variables. After we decompose the structure, we will finally get the results given by Bouchaud & Potters [1] on the moments.

In chapter 3, we deal with the resolvent matrix. which is defined by

$$\left\langle \frac{1}{z - E} \right\rangle.$$

It contains more information than the Stieltjes transform, and the Stieltjes transform is just the trance of the resolvent matrix. In Sengupta & Mitra's paper [2], they gave a result on the resolvent matrix, equation (30). In their paper, they mentioned a diagrammatic method from quantum field theory to deal with the resolvent matrix. However, they did not provide any details about this method. In Burda et al.'s papers, [3][4], they also mentioned this method and gave more details. However, the derivation still relies heavily on the arguments from quantum field theory. We give a more detailed treatment to the diagrammatic method, to make the approach understandable to anyone with a non-physics background. In the end of chapter 3, we derive a recurrence formula based on Sengupta & Mitra's result, and we shall see this gives the same result as Bouchaud & Potters's result.

In chapter 4, we will extend the diagrammatic method and apply it to compute the moments of E in terms of C and D. We translate the problem of computing the moments to the problem of counting the non-crossing partitions of the set $[k]$, for some $k \in \mathbb{N}$. In order to solve the counting problems, we will borrow some tools from combinatorics. We are going to establish several one-to-one correspondences, and finally convert the counting problem to a solved problem of representation theory, and its solution is given by Goulden & Jackson [5].

Chapter 2

Free Probability Theory

In classical probability theory, a random variable is a measurable function

$$f\colon \Omega \longrightarrow \mathbb{K}$$

from the measure space $(\Omega, \mathcal{A}, P)$ to the field $\mathbb{K} = \mathbb{R}$ or $\mathbb{C}$. We could add or multiply any two random variables, this is due to the ring structure of $\mathbb{K}$. More specifically, let $L^1(\mathbb{K})$ be the set of all random variables with finite expectation, then $L^1(\mathbb{K})$ is a $\mathbb{K}$-algebra. The properties of space of functions usually depend on the properties of the target space. Another important property of $L^1(\mathbb{K})$ comes from $\mathbb{K}$ is the commutativity of multiplication. Suppose we change the field $\mathbb{K}$ to some matrix spaces, for example, the general linear group $\mathrm{GL}(k)$. $\mathrm{GL}(k)$ has natural topology structure, and hence, it can induce a Borel measure on $\mathrm{GL}(k)$ and then we could define $\mathrm{GL}(k)$-valued the measurable function on Ω. This kind of random variables are not commutative under multiplication due to the non-commutativity of $\mathrm{GL}(k)$.

Free probability theory was initiated by Dan Voiculescu [6]. It is a more general theory of non-commutative probability, which does not have the underlying event space Ω. Many basic notions can be discussed in a purely algebraic context. The idea freeness from free probability theory is the analogy to the independence from classical probability theory, which will be very useful to us to study the product in (1.3).

In this chapter, we give a quick introduction to the free probability theory and then use it to derive the results achieved by Bouchaud & Potters on moments in [1][7].

2.1 Non-Commutative Probability Space

We first treat some basic operator algebras.

Definition 2.1.1. *A $*$-algebra $(\mathcal{A}, *)$ consists of a $\mathbb{C}$-algebra $\mathcal{A}$ and an operator $*\colon \mathcal{A} \longrightarrow \mathcal{A}$ such that for all $X, Y \in \mathcal{A}$ and $c \in \mathbb{C}$*

1. $(X+Y)^* = X^* + Y^*$;
2. $(XY)^* = Y^*X^*$;
3. $1^* = 1$;
4. $X^{**} = X$;
5. $(cX)^* = \overline{c}X^*$.

A linear functional $\tau\colon \mathcal{A} \longrightarrow \mathbb{C}$ is $*$-linear if for every $x \in \mathcal{A}$, we have $\tau(A^*) = \overline{\tau(A)}$. A non-commutative probability space $(\mathcal{A}, *, \tau)$ consists of a $*$-algebra $(\mathcal{A}, *)$ and a $*$-linear functional τ such that for every $X, Y \in \mathcal{A}$, we have

1. $\tau(1) = 1$;
2. $\tau(X^*X) \leq 0$; *(Non-negativity)*

3. $\tau(XY) = \tau(YX)$; *(Trace axiom)*

4. $\tau(X^*X) = 0$ *if and only if* $X = 0$. *(Faithfulness)*

We often just write $(\mathcal{A}, \tau)$ *for a non-commutative probability space. The elements in* $\mathcal{A}$ *are called random variables,* τ *is called trace.*

Generally, the last two axioms of non-commutative probability space could be removed, but all the examples that we are going to consider satisfy all these axioms.

Example 2.1.1. *(Classical random variables) Consider a sample space* Ω *we use* $L^{\infty-}$ *to denote the space of all random variables defined on* Ω *such that all their moments are finite. This is actually a very strong condition, but most of the classical random variables that we are going to consider are Gaussian or derived from Gaussian, thus, this condition is satisfied. Complex conjugation plays the role of the* $*$*-operator and the usual expectation* $\mathbb{E}(\cdot)$ *plays the role of trace. Hence* $(L^{\infty-}, \mathbb{E})$ *is a non-commutative probability space, even though it is actually commutative.*

Example 2.1.2. *(Deterministic matrices) Consider the space of all* $N \times N$ *complex valued matrices denoted by* $M_N(\mathbb{C})$*. It is a* $*$*-algebra, where the* $*$*-operator is given by the matrix conjugate transpose,* $X^* = \overline{X^T}$*. We use normalized trace defined by* $\tau(X) := \frac{1}{N}\mathrm{tr}\,(X)$ *as the trace operator. The reason to normalize it is we want it maps the identity matrix* I_N *to* 1*. It is easy to check that* $\frac{1}{N}\mathrm{tr}\,(\cdot)$ *satisfies all axioms to become a trace. Hence,* $(M_N(\mathbb{C}), \tau)$ *forms a non-commutative probability space.*

Example 2.1.3. *(Random matrices) Consider the space* $L^{\infty-} \otimes M_N(\mathbb{C})$*, this could be thought as a* $N \times N$ *matrix with all its entries are random variables from* $L^{\infty-}$*. It is a* $\mathbb{C}$*-algebra, its structure is induced by the tensor product of two* $\mathbb{C}$*-algebras. As a result, both* $L^{\infty-}$ *and* $M_N(\mathbb{C})$ *have natrual monomorohisms to* $L^{\infty-} \otimes M_N(\mathbb{C})$ *given by* $L^{\infty-} \ni X \mapsto X \otimes I_N$ *and* $M_N(\mathbb{C}) \ni A \mapsto \mathbb{1} \otimes A$*. Since the multiplicative identity of* $L^{\infty-} \otimes M_N(\mathbb{C})$ *is* $\mathbb{1} \otimes I_N$*, the trace operator is given by*

$$\tau(X) := \mathbb{E}\frac{1}{N}\mathrm{tr}\,(X) = \frac{1}{N}\mathrm{tr}\,(\langle X \rangle).$$

$\langle \cdot \rangle$ *means the empirical average of the realization matrices of* X*, could be thought as taking the expectation to all the entries of the random matrix* X*.* $\langle \cdot \rangle$ *is the same as* $\mathbb{E}(\cdot)$ *when we apply it to the classical random variables. We use the same letter* τ *as the trace operator from the* M_n *case, this is because two trace operators agree on the image of* $M_N(\mathbb{C})$ *under the natural monomorphism mentioned above. One can check that* τ *satisfies all the axioms and then* $(L^{\infty-} \otimes M_N(\mathbb{C}), \tau)$ *is a non-commutative probability space.*

2.2 Moments and Spectrum

Definition 2.2.1. *For a random variable* $X \in \mathcal{A}$*, we define its* k*-th moment to be* $\tau(X^k)$ *for any* $k \in \mathbb{N}$*.*

In classical probability theory, we could compute the k-th-moment of a random variables X though

$$\tau(X^k) = \int_{\mathbb{C}} z^k d\mu_X,$$

where μ_X is the probability measure of X on $\mathbb{C}$. For a deterministic matrix A, we assume it is self-adjoint, that is $A^* = A$. Hence, A is diagonalizable and we have

$$\tau(A^k) = \frac{1}{N}\mathrm{tr}\,(A^k) = \sum_{i=1}^{N} \lambda_i^k,$$

where λs are the eigenvalues of A and are all real. We can also write the above sum into

$$\int_{-\infty}^{+\infty} x^k dF_A,$$

where

$$F_A(x) = \frac{1}{N}\sum_{i=1}^{N} \mathbb{1}_{\{x \le \lambda_i\}}.$$

Note that F_A could be thought as the cumulative distribution function of some probability measure μ_A on $\mathbb{C}$, more specifically, on the real line. This measure is called the spectrum of A.

Now, as for the random matrix case, we want to define the spectrum of a random matrix $A \in L^{\infty-} \otimes M_n(\mathbb{C})$ such that A is self-adjoint. By analogy to the previous case, we are looking for a probability measure μ_A such that for any $k \in \mathbb{N}$

$$\tau(A^k) = \frac{1}{N}\mathbb{E}\mathrm{tr}\,(A^k) = \int_{-\infty}^{+\infty} x^k d\mu_A.$$

We integrate it over the real line since A is self-adjoint, or equivalently, hermitian. Hence, its eigenvalues are real-valued random variables.

Consider a polynomial $p \in \mathbb{C}[z]$, then $\tau(p(A))$ defines a linear functional on $\mathbb{C}[z]$. This functional is actully bounded on the closed interval $[-\rho(A), \rho(A)]$ where $\rho(A) := \|\|A\|_{op}\|_{L^\infty}$, the essential norm. The detail of this part could be found in [8], section 2.5.1. The boundedness allows us to use the Weierstrass approximation theorem. The bounded linear functional $p \mapsto \tau(p(A))$ can be extended to $C_c([-\rho(A), \rho(A)])$, the space consists of all continuous functions defined on $[-\rho(A), \rho(A)]$. Therefore, according to Riesz–Markov–Kakutani representation theorem, there exists a unique Radon measure μ_A such that

$$\int_{-\infty}^{+\infty} \varphi(x) d\mu_A = \mathbb{E}\int_{-\infty}^{+\infty} \varphi(x) d\mu_M,$$

where M are the realizations of A and is deterministic. Since we have $\tau(I_N) = 1$, μ_A is a probability measure. We call such measure the spectrum of A. This result has a general version for the self-adjoint random variables of an arbitrary non-commutative space, see [8] for the details.

2.3 Stieltjes Transform

In order to compute μ_A for a hermitian random matrix, we need to use the Stieltjes transform.

Definition 2.3.1. *The Stieltjes transform of a probability measure μ on $\mathbb{R}$ is a complex-valued function defined by*

$$\mathfrak{g}_\mu(z) := \int_{-\infty}^{+\infty} \frac{1}{z-x} d\mu(x)$$

Note that $\mathfrak{g}_\mu$ is regular at all points that is not a support of μ, in particular, $\mathfrak{g}_\mu$ is regular on $\mathbb{C} \setminus \mathbb{R}$.

If μ comes from a random variable A and the support of μ_A is bounded, then $\mathfrak{g}_{\mu_A}(z)$, or we can write as $\mathfrak{g}_A(z)$, is regular at infinity, and hence outside a disc of sufficiently large radius, we can expand

$$\mathfrak{g}_A(z) = \int_{-\infty}^{+\infty} \sum_{k=0}^{\infty} \frac{x^k}{z^{k+1}} d\mu(x) = \sum_{k=0}^{\infty} \frac{\tau(A^k)}{z^{k+1}} = \tau\left(\frac{1}{z-A}\right). \tag{2.1}$$

We define the resolvent matrix of the hermitian matrix A to be

$$\mathcal{G}_A(z) := \frac{1}{z-A}. \tag{2.2}$$

More precisely, since A is hermitian, we can write $A = O\Lambda O^*$ where Λ is real-valued diagonal matrix and we have $OO^* = I$. Therefore,

$$\mathcal{G}_A(z) = \frac{1}{z-A} = O\left(\sum_{k=0}^{\infty}(\frac{\Lambda^k}{z^{k+1}})\right)O^T$$

is well-defined. Note that $\mathfrak{g}_A = \tau(\mathcal{G}_A)$.

The Stieltjes transform has inverse formulae, we could recover the measure by the inverse Stieltjes transform. The following formula is called Sokhotski–Plemelj folmula, we could use it to recover the density function $f_\mu(x)$ of μ

$$f_\mu(x) = \frac{1}{\pi}\lim_{b\to 0^+} \mathrm{Im}\,\mathfrak{g}(x-ib). \tag{2.3}$$

The derivation of this formula could be found in [1] or in [8]. If one does not comfortable with Dirac measure, for a discrete measure we could recover its cumulative distribution function $F_\mu(x)$ by

$$F_\mu(x) = \frac{1}{\pi} \lim_{b \to 0^+} \int_{-\infty}^{x} \operatorname{Im} \mathfrak{g}(a - ib) da. \tag{2.4}$$

The derivation of this formula could be found in [9], chapter 13. Note that the existence of inverse fomula suggests that the if we know all the moments of a random variables, then we know its spectrum.

Now let A be a $N \times N$ hermitian random matrix, with eigenvalues $\lambda_1, \ldots, \lambda_N$. Note that λs are real-valued classical random variables. Then we have

$$\begin{aligned} \mathfrak{g}_A(z) &= \sum_{k=0}^{\infty} \frac{\tau(A^k)}{z^{k+1}} \\ &= \sum_{k=0}^{\infty} \frac{\mathbb{E}(\sum_{i=1}^{N} \lambda_i^k)}{N z^{k+1}} \\ &= \frac{1}{N} \mathbb{E} \left(\sum_{k=0}^{\infty} \frac{\sum_{i=1}^{N} \lambda_i^k}{z^{k+1}} \right) \\ &= \frac{1}{N} \mathbb{E} \left(\sum_{i=1}^{N} \frac{1}{z - \lambda_i} \right). \end{aligned} \tag{2.5}$$

By equation (2.4), we get

$$\begin{aligned} F_A(x) &= \frac{1}{\pi} \lim_{b \to 0^+} \int_{-\infty}^{x} \operatorname{Im} \frac{1}{N} \mathbb{E} \left(\sum_{i=1}^{N} \frac{1}{a - ib - \lambda_i} \right) da \\ &= \frac{1}{N\pi} \mathbb{E} \lim_{b \to 0^+} \int_{-\infty}^{x} \sum_{i=1}^{N} \frac{b}{(a - \lambda_i)^2 + b^2} da \\ &= \frac{1}{N\pi} \mathbb{E} \lim_{b \to 0^+} \sum_{i=1}^{N} \left(- \arctan \frac{\lambda_i - x}{b} + \frac{\pi}{2} \right) \\ &= \mathbb{E} \left(\frac{1}{N} \sum_{i=1}^{N} \mathbb{1}_{\{x \leq \lambda_i\}} \right). \end{aligned} \tag{2.6}$$

We have a very important theorem on Stieltjes transform, that is the Stieltjes continuity theorem

Theorem 2.3.2. *(Stieltjes continuity theorem) Let μ_n be a sequence of probability measure on $\mathbb{R}$ and μ_∞ an additional probability measure on $\mathbb{R}$. Then μ_n converges to μ_∞ weakly if and only if $\lim_{n \to \infty} \mathfrak{g}_{\mu_n}(z) = \mathfrak{g}_{\mu_\infty}(z)$ for all $z \in \mathbb{C}$ with positive imaginary part.*

Proof. Omitted. See [10], Theorem 4.14. □

This theorem would be very useful when we need to discuss the large random matrices.

2.4 Limit and Marčenko-Pastur Distribution

Definition 2.4.1. *Suppose we have a sequence of non-commutative probability space $(\mathcal{A}_n, \tau_n)$ and an additional space $(\mathcal{A}_\infty, \tau_\infty)$. We say that a sequence of random variables $X_n \in \mathcal{A}_n$ converges in the sense of the moments to a random variable $X_\infty \in \mathcal{A}_\infty$ if for any $k \in \mathbb{N}$, we have*

$$\tau(X_n^k) \longrightarrow \tau(X_\infty^k).$$

Note that the convergence of arbitrary moment implies that $\mathfrak{g}_{X_n(z)}$ converges to $\mathfrak{g}_{X_\infty}(z)$ for $|z|$ sufficiently large, then by identitiy theorem, we could conclude that $\mathfrak{g}_{X_n(z)}$ converges to $\mathfrak{g}_{X_\infty}(z)$

for all z that is regular. Now suppose X_n and X_∞ are all self-adjoint, then by Stieltjes continuity theorem, convergence in the sense of moment of X_n is equivalent to convergence of their Stieltjes transform and is also equivalent to the weak convergence of the measures induced by X_n.

By this construction, we are able to talk about the asymptotic behaviour of a random matrix when its dimension is large. We had seen from equation (2.6), the spectrum of a finite dimensional hermitian random matrix is discrete. But we could take the limit of its Stieltjes transform and then use inverse Stieltjes transform to get, ideally, a continuous density.

Recall that the empirical covariance matrix is given by $E = \frac{XX^T}{T}$, where X is $N \times T$ real-valued random matrix satisfies each its column and row are Gaussian. Since the (i, j) entry of E, $E_{ij} = T^{-1} \sum_t X_{it} X_{jt} = E_{ji}$, hence E is an symmetric random matrix. In [11], Marčenko and Pastur gives a result on the spectrum of E when the true covariance matrices C and D are all identity matrices, then equation (1.2) becomes

$$\mathbb{E}(X_{it} X_{js}) = \delta_{ij}.$$

When taking the limit, we also assume that the ratio N/T stays as a constant and denoted it by q.

Theorem 2.4.2. *(Marčenko-Pastur law) The limit of Stieltjes transform $\mathfrak{g}(z)$ of empirical covariance matrix satisfies*

$$\frac{1}{\mathfrak{g}(z)} = z - 1 + q - qz\mathfrak{g}(z), \tag{2.7}$$

where $q = N/T < 1$ is a constant. Hence, we could solve $\mathfrak{g}(z)$ and get

$$\mathfrak{g}(z) = \frac{z - (1 - q) - \sqrt{z - \lambda_+}\sqrt{z - \lambda_-}}{2qz} \tag{2.8}$$

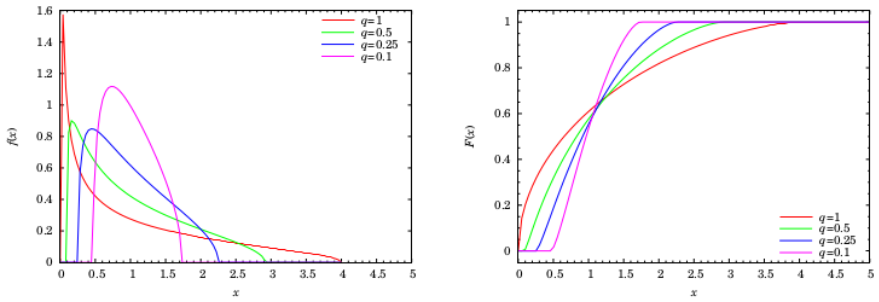
for all $z \notin (\lambda_-, \lambda_+)$ where

$$\lambda_\pm = (1 \pm \sqrt{q})^2.$$

Furthermore, by inverse Stieltjes transform, we can get its density

$$f(x) = \frac{\sqrt{(\lambda_+ - x)(x - \lambda_-)}}{2\pi q x} \qquad \text{where } \lambda_- < x < \lambda_+ \tag{2.9}$$

A complete derivation of this theorem could be found in [1]. Note that equation (2.8) is not the only solution to the equation (2.7), which branches to choose is also discussed in [1], section 4.2.3.



(i) (ii)

Figure 2.1: (i) Marčenko–Pastur density for $q = 1, 0.5, 0.25, 0.1$. (ii) Cumulative distribution function, for these values of 1, and estimate obtained empirically, for three different realisations of the underlying matrix. In each case $N = 100$.

2.5 Freeness

Now we want to deal with the equation (1.3), that is

$$E = \frac{XX^T}{T} = \frac{C^{\frac{1}{2}} X_0 D X_0^T C^{\frac{1}{2}}}{T}$$

where the entries of X_0 are independent Gaussian variables. According to Marčenko-Pastur law, we know how to deal with $X_0 X_0^T$. To understand the product above, we need the concept of freeness.

Definition 2.5.1. *Two random variables $X, Y \in \mathcal{A}$ are free if for any finite collection of polynomials $p_1, \ldots, p_n, q_1, \ldots, q_n \in \mathbb{C}[x]$ satisfies for any i, $\tau(p_i(X)) = 0$ and $\tau(q_i(Y)) = 0$, we have*

$$\tau\left(\prod_i p_i(X) q_i(Y)\right) = 0.$$

Two sequences of random variables X_n and Y_n are asymptotically free if for any finite collection of polynomials $p_1, \ldots, p_n, q_1, \ldots, q_n \in \mathbb{C}[x]$ satisfies for any i, $\tau(p_i(X_n)) = 0$ and $\tau(q_i(Y_n)) = 0$, we have

$$\tau\left(\prod_i p_i(X_n) q_i(Y_n)\right) \longrightarrow 0$$

as $n \longrightarrow \infty$.

Suppose we have two sequences of real symmetric random matrices A_n and B_n, where A_n, B_n are $n \times n$ then we could write A_n, B_n into

$$A_n = U_n \Lambda_n U_n^T, \qquad B_n = V_n \Lambda_n' V_n^T,$$

where U_n, V_n are $n \times n$ orthogonal random matrices, and Λ_n, Λ_n' are $n \times n$ diagonal random matrices. Then for any collection of traceless polynomials p_i, q_i

$$\tau\left(\prod_i p_i(A_n) q_i(B_n)\right) = \tau\left(\prod_i p_i(\Lambda_n) U_n^T V_n q_i(\Lambda_n') V_n^T U_n\right).$$

By observing this, we notice that if the sequence of random matrices $U_n^T V_n$ tends to zero matrix, then A_n, B_n are asymptotically free.

In order to apply this result to empirical covariance matrix E, we need some knowledge of Haar measure.

Definition 2.5.2. *A topological group G is a topological space that is also a group at the same time. The group operation and inverse operation are continuous under the topology it is equipped with.*

A measure H on a topological group G is called the Haar measure over the Borel sigma algebra, if $H(gJ) = H(J)$ for any $g \in G$ and $J \subseteq \mathcal{B}(G)$.

Every locally compact topological group has a finite Haar measure, and such measure is unique up to multiplication by a constant, details can be found in [12], chapter 9. From now on, we assume the Haar measure H satisfies $H(G) = 1$, then H defines a probability measure.

The group of orthogonal matrices $O(n)$ is a typical locally compact topological group, hence, it has a unique Haar measure H.

Proposition 2.5.3. *Suppose $X \in O(n)$ is Haar distributed, then for any deterministic vector $v \in S^{n-1}$, Xv is uniformly distributed on the sphere S^{n-1}.*

Proof is Omitted, see [13], chapter 10.

Two vectors drawn uniformly from S^n are orthogonal almost surely as n tends to infinity. In fact, the inner product of any pair of distinct vectors with unit norm tends to zero as the dimension tends to zero. In [13], they had also argued that the eigenmatrix of $E = \frac{X_0 X_0^T}{T}$ is Haar distributed. Therefore, the white E is free to any deterministic matrix, and any two deterministic matrices are also free.

2.6 Results Given by Bouchaud & Potters

Definition 2.6.1. *Let A be a random matrix, we define the T-transform to be*

$$\mathcal{T}_A(z) = \sum_{k=1}^{\infty} \frac{\tau(A^k)}{z^k} = z\mathfrak{g}_A(z) - 1.$$

Note that $\mathcal{T}$ is invertible for large z if A is not traceless. Hence $\mathcal{T}^{-1}(t)$ exists for small t. Hence, this allows us to define:

Definition 2.6.2. *We define the S-transform to be*

$$\mathcal{S}_A(t) = \frac{t+1}{t\mathcal{T}_A^{-1}(z)}.$$

Let us see an important example.

Example 2.6.1. *By equation (2.7), we have*

$$\frac{1}{\mathfrak{g}(z)} = z - 1 + q - qz\mathfrak{g}(z).$$

By substituting $\frac{\mathcal{T}(z)+1}{z}$ for $\mathfrak{g}(z)$, we get

$$\begin{aligned} \frac{z}{\mathcal{T}(z)+1} &= z - 1 - q\mathcal{T}(z) \\ z\mathcal{T}(z) &= (1 + q\mathcal{T}(z))(\mathcal{T}(z) + 1) \end{aligned}$$

By substituting $\mathcal{T}^{-1}(t)$ for z, we have

$$\begin{aligned} \mathcal{T}^{-1}(t)\mathcal{T}(\mathcal{T}^{-1}(t)) &= (1 + q\mathcal{T}(\mathcal{T}^{-1}(t)))(\mathcal{T}(\mathcal{T}^{-1}(t)) + 1) \\ \mathcal{T}^{-1}(t) &= (1 + qt)(t + 1) \end{aligned}$$

By rearranging, we finally get

$$\mathcal{S}_{T^{-1}X_0X_0^T}(t) = \frac{t+1}{t\mathcal{T}^{-1}(t)} = \frac{1}{1+qt} \tag{2.10}$$

The most important property of S-transform is the following:

Theorem 2.6.3. *Let A, B be two free random variables, then we have*

$$\mathcal{S}_{AB}(t) = \mathcal{S}_A(t)\mathcal{S}_B(t). \tag{2.11}$$

We omit the proof, one can found a complete proof in [1].

This multiplicative property could be passed to T-transform. According to equation (2.11), we have

$$\frac{t+1}{t\mathcal{T}_{AB}^{-1}(t)} = \frac{t+1}{t\mathcal{T}_A^{-1}(t)}\mathcal{S}_B$$

which implies

$$\mathcal{T}_{AB}^{-1}(t)\mathcal{S}_B = \mathcal{T}_A^{-1}(t). \tag{2.12}$$

Then we have

$$\begin{aligned} \mathcal{T}_{AB}(z) &= \mathcal{T}_A\left(\mathcal{T}_A^{-1}(\mathcal{T}_{AB}(z))\right) \\ &= \mathcal{T}_A\left(\mathcal{T}_{AB}^{-1}(\mathcal{T}_{AB}(z))\mathcal{S}_B(\mathcal{T}_{AB}(z))\right) \\ &= \mathcal{T}_A\left(z\mathcal{S}_B(\mathcal{T}_{AB}(z))\right) \end{aligned} \tag{2.13}$$

Now we are ready to deal with the empirical covariance matrix $E = T^{-1}C^{\frac{1}{2}}X_0DX_0^TC^{\frac{1}{2}}$ as we had discussed. By trace axiom, we have $\mathcal{T}_E(z) = \mathcal{T}_{T^{-1}CX_0DX_0^T}(z)$, and hence true for S-transform, then by freeness, we get

$$\mathcal{S}_E(t) = \mathcal{S}_C(t)\mathcal{S}_{T^{-1}X_0DX_0^T}(t)$$

Note that

$$\frac{1}{N}\mathrm{tr}\left(\left(\frac{X_0DX_0^T}{T}\right)^k\right) = \frac{1}{qT}\mathrm{tr}\left(\left(\frac{DX_0^TX_0}{T}\right)^k\right).$$

This can be concluded by

$$\mathcal{T}_{T^{-1}X_0DX_0^T}(z) = \frac{1}{q}\mathcal{T}_{T^{-1}DX_0^TX_0}(z).$$

By taking the inverse, we get

$$\mathcal{T}^{-1}_{T^{-1}X_0DX_0^T}(t) = \mathcal{T}^{-1}_{T^{-1}DX_0^TX_0}(qt).$$

By definition of S-transform, we have

$$\begin{aligned}
\mathcal{S}_{T^{-1}X_0DX_0^T}(t) &= \frac{t+1}{t\mathcal{T}^{-1}_{T^{-1}DX_0^TX_0}(qt)} \\
&= \frac{qt+q}{qt}\frac{qt+1}{qt+1}\frac{1}{\mathcal{T}^{-1}_{T^{-1}DX_0^TX_0}(qt)} \\
&= \frac{qt+q}{qt+1}\mathcal{S}_{T^{-1}DX_0^TX_0}(qt) \\
&= \frac{qt+q}{qt+1}\mathcal{S}_D(qt)\mathcal{S}_{T^{-1}X_0^TX_0}(qt).
\end{aligned}$$

By similar argument as Example 2.6.1, we can get $\mathcal{S}_{T^{-1}X_0^TX_0}(qt) = \frac{1}{qt+q}$. Finally, we derive

$$\mathcal{S}_E(t) = \frac{\mathcal{S}_C(t)\mathcal{S}_D(qt)}{1+qt}. \tag{2.14}$$

We had seen that by using S-transform, we can decompose the structure of equation (1.3). By expanding equation (2.14), we have

$$\begin{aligned}
\frac{t+1}{t\mathcal{T}_E^{-1}(t)} &= \frac{t+1}{t\mathcal{T}_C^{-1}(t)}\frac{1}{qt\mathcal{T}_D^{-1}(qt)} \\
\mathcal{T}_E^{-1}(t) &= qt\mathcal{T}_D^{-1}(t)\mathcal{T}_C^{-1}(t)
\end{aligned}$$

$$\begin{aligned}
\mathcal{T}_E^{-1}(\mathcal{T}_E(z)) &= q\mathcal{T}_E(z)\mathcal{T}_D^{-1}(\mathcal{T}_E(z))\mathcal{T}_C^{-1}(\mathcal{T}_E(z)) \\
q\mathcal{T}_E(z) &= \mathcal{T}_D\left(\frac{z}{q\mathcal{T}_E(z)\mathcal{T}_C^{-1}(\mathcal{T}_E(z))}\right).
\end{aligned} \tag{2.15}$$

This result is given by Bouchaud & Potters in [1], equation (17.44). By expanding the left hand side into Laurent series and compare the coefficient on both side, we could compute the moments of E in terms of the moment of C. But this is very difficult to do this, since this formula involving the computation of functional inverse of a Laurent series.

The special case when D is trivial would be much easier. One can check that the S-transform of the identity matrix is $\frac{1}{z-1}$. Then by rearranging equation (2.15), we will get

$$\mathcal{T}_E(z) = \mathcal{T}_C\left(\frac{z}{1+q\mathcal{T}_E(z)}\right).$$

In the language of Stieltjes Transform,

$$z\mathfrak{g}_E(z) = Z\mathfrak{g}_C(Z), \quad \text{where } Z = \frac{z}{1-q+qz\mathfrak{g}_E(z)}. \tag{2.16}$$

This is one of the main result of Bouchaud & Potters [7][1]. We could expand it and get the equation (17.10) from [1].

$$\sum_{k=1}^{\infty} \frac{\tau(E^k)}{z^k} = \sum_{k=1}^{\infty} \frac{\tau(C^k)}{z^k} \left(1 + q \sum_{l=1}^{\infty} \frac{\tau(E^l)}{z^l}\right)^k. \tag{2.17}$$

This can be used to compute the moments of E by comparing the coefficients, and we get

$$\begin{aligned} \tau(E^1) &= \tau(C) \\ \tau(E^2) &= \tau(C^2) + q\tau(C)^2 \\ \tau(E^3) &= \tau(C^3) + 3q\tau(C^2)\tau(C) + q^2\tau(C)^3. \end{aligned}$$

But this method still be very slow, since the calculation involve the computation of functional composition of power series.

Recall that the resolvent of the empirical covariance matrix should be

$$\mathcal{G}_E(z) = \left\langle \frac{I_N}{zI_N - T^{-1}XX^T} \right\rangle = \sum_{k=0}^{\infty} \frac{\langle E^k \rangle}{z^{k+1}}, \qquad \text{for } |z| \text{ sufficient large,} \tag{2.18}$$

Suppose the equation (2.16) also holds for the resolvent matrix[1], that is

$$z\mathcal{G}_E(z) = Z\mathcal{G}_C(Z), \quad \text{where } Z = \frac{z}{1 - q + qz\mathfrak{g}_E(z)}. \tag{2.19}$$

We can rewrite it and get

$$z\mathcal{G}_C(z) - I = \big(1 - q + qz\mathfrak{g}_E(z)\big)C\mathcal{G}_E(z). \tag{2.20}$$

If we write $\tau(E^k)$ as G_k, we will achieve a self-convolutive recurrence: expanding the above and equating terms in z^{-k-1},

$$G_{k+1} = CG_k + q\sum_{j=1}^{k} \tau(G_j)CG_{k-j}, \qquad k \geq 0. \tag{2.21}$$

This gives a more clear form to see how the moment is compute, and we will use this to compare the other methods that we are going to mention later.

[1]This is indeed true, from private communication with Bouchaud

Chapter 3

Diagrammatic Method

We had seen the Bouchaud & Potters' result on the moments and the resolvent. There is another result on the resolvent matrix derived in Sengupta & Mitra's paper [2].

In their paper, they using the techniques from physics and the Feynman diagrams, and found that $\mathcal{G}_E$ should satisfies

$$\mathcal{G}_E(z) = \frac{I_N}{zI_N - C\mathrm{tr}\left(\dfrac{D}{I_T - D\mathrm{tr}\left(C\mathcal{G}_E(z)\right)}\right)}. \tag{3.1}$$

Unfortunately, Sengupta & Mitra did not give the details of this method. In Burda et al.'s paper [3][4], they gives more details about this method, and give the following equations

$$\begin{aligned} \mathcal{G}_E(z) &= \frac{I_N}{zI_N - \Sigma(z)}, \qquad \Sigma(z) = C\mathrm{tr}\left(D\mathcal{G}_{E*}(z)\right), \\ \mathcal{G}_{E*}(z) &= \frac{I_T}{TI_T - \Sigma_*(z)}, \quad \Sigma_*(z) = D\mathrm{tr}\left(C\mathcal{G}_E(z)\right). \end{aligned} \tag{3.2}$$

$\mathcal{G}_{E*}$ in above is the diagrammatic dual of $\mathcal{G}_E$, which we will explain it later, it is defined by

$$\mathcal{G}_{E*} := \left\langle \frac{I_T}{TI_T - z^{-1}X^TX} \right\rangle. \tag{3.3}$$

In these papers, however, they still omit quite a few details, which can be difficult for readers who are unfamiliar with the quantum field theory.

In this chapter, we use the same diagrammatic setting as the one in [3], and using completely combinatorial and mathematical language to derive this result.

3.1 Diagrammatic Structure for Resolvent

Firstly, by linear algebra, for an arbitrary $k \in \mathbb{N}$, the (i,j) entry of $\left\langle E^k \right\rangle$ is

$$\left\langle E^k \right\rangle_{ij} = \sum_{(i_1,\ldots,i_{k-1})\in[N]^{k-1}} \left\langle E_{ii_1}E_{i_1i_2}\cdots E_{i_{k-2}i_{k-1}}E_{i_{k-1}j} \right\rangle. \tag{3.4}$$

Similarly, we have

$$(C^k)_{ij} = \sum_{(i_1,\ldots,i_{k-1})\in[N]^{k-1}} C_{ii_1}C_{i_1i_2}\cdots C_{i_{k-2}i_{k-1}}C_{i_{k-1}j}. \tag{3.5}$$

Note that (3.4) and (3.5) have similar summing terms. How these terms related depends on the specific probability measure. Since we had assumed that the entries of X are Gaussian distributed, we can study the expectation though Wick's theorem [1].

Theorem 3.1.1. *(Wick's theorem) Suppose $X_1, \ldots, X_{2n}$ are zero mean multivariate Gaussian random variables. Then we have*

$$\langle X_1 \ldots X_{2n} \rangle = \sum_{parings} \prod_{pairs} \langle X_i X_j \rangle , \tag{3.6}$$

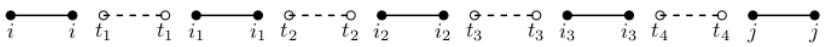
where pairing pairs means summing over all the possible combinations of distinct pairings in the form of product.

By applying Wick's theorem to equation (3.4), we can expand

$$\begin{aligned} \left\langle E^k \right\rangle_{ij} &= \frac{1}{T^k} \sum_{\substack{(i_1,\ldots,i_{k-1}) \in [N]^{k-1} \\ (t_1,\ldots,t_k) \in [T]^k}} \left\langle X_{it_1} X_{i_1 t_1} X_{i_1 t_2} X_{i_2 t_2} \cdots X_{i_{k-1} t_k} X_{j t_k} \right\rangle \\ &= \frac{1}{T^k} \sum_{\substack{(i_1,\ldots,i_{k-1}) \in [N]^{k-1} \\ (t_1,\ldots,t_k) \in [T]^k}} \sum_{parings} \prod_{pairs} \cdot \langle X_{\alpha\beta} X_{\gamma\delta} \rangle \end{aligned}$$

The diagrammatic method is a way to enumerate these pairing pairs, and these pairing pairs could be grouped into a "non-planar" group and "planar" group, which they have different behaviours when N, T tends to infinity with N/T remains as a constant q.

Diagram setting We use black circles and white circles to represent the spatial indices and temporal indices respectively. We use $\bullet\,\circ$ to represent an element from X and $\circ\,\bullet$ to represent an element from X^T. We use solid lines and dashed lines to connect two circles that have the same spatial and temporal index respectively. To make the diagram complete, we add two end points to the diagram that are currently meaningless. This notation would be very helpful to counting the trace in the sum and study the structure. For example, if $k = 4$, we have



Now we use solid arcs and dashed arcs to pair these Xs. If we want to pair X_{it} with X_{js}, we use a solid arc to connect their spatial indices, and we use a dashed arc to connect their temporal indices. For example,

$$\langle X_{it_1} X_{i_3 t_3} \rangle \langle X_{i_1 t_1} X_{i_1 t_2} \rangle \langle X_{i_2 t_2} X_{i_2 t_3} \rangle \langle X_{i_3 t_4} X_{j t_4} \rangle \tag{3.7}$$

can be expressed by the following diagram

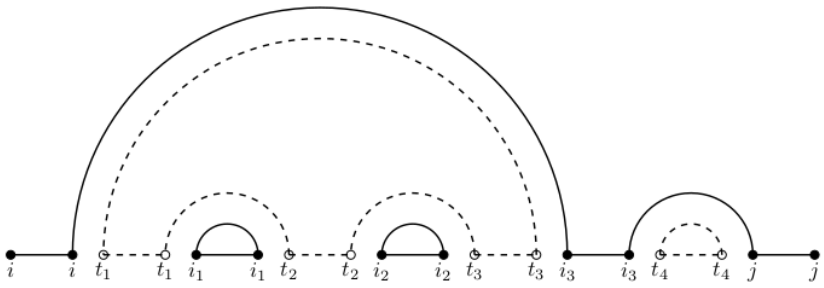


Figure 3.1: The diagram representation of equation (3.7)

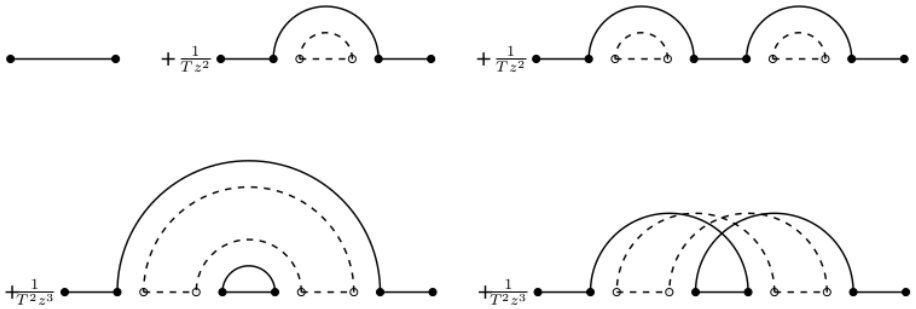
By our assumption $\langle X_{\alpha\beta} X_{\gamma\delta} \rangle = C_{\alpha\gamma} D_{\beta\delta}$, a solid arc will give us a C_{ij} and a dashed line will give us an D_{ts}. The above diagram give us

$$C_{ii_3} D_{t_1 t_3} C_{i_1 i_1} D_{t_1 t_2} C_{i_2 i_2} D_{t_2 t_3} C_{i_3 j} D_{t_4 t_4}. \tag{3.8}$$

Now we remove all the indices on the diagram to denote the sum runs over all the possible indices with fixed i, j. Then we have

$$\begin{aligned}\left\langle E^k \right\rangle_{ij} &= \sum_{\substack{i_1,\cdots i_{k-1}\\ t_1,\cdots t_k}} \sum_{\substack{indexed\\ diagrams}} \\ &= \sum_{\substack{indexed\\ diagrams}} \sum_{\substack{i_1,\cdots i_{k-1}\\ t_1,\cdots t_k}} \\ &= \sum_{\substack{indexless\\ diagrams}}\end{aligned}$$

This allows us to write $(\mathcal{G}_E)_{ij}$ as the following series



Note that the power of z equals to the number of solid horizontal lines and the power of T equals to the number of dashed horizontal lines. So we can omit the coefficient in the front of the diagrams, we can recover it as we like.

3.2 Structure of a Single Diagram

In Figure 3.1, one dashed-loop could be given by t_1, t_2, t_3, t_1. It corresponds to the product $D_{t_1t_2}D_{t_2t_3}D_{t_3t_1}$ from equation (3.8). This allow us to write our expression in the form of trace. If we summing equation (3.8) over all possible indices and take the coefficient in $\mathcal{G}_E$ into account, we get

$$\begin{aligned}&\frac{1}{T^4}\sum (C_{ii_3}C_{i_3j})(D_{t_1t_3}D_{t_1t_2}D_{t_2t_3})C_{i_1i_1}C_{i_2i_2}D_{t_4t_4}. \\ =\ &\frac{1}{T^4}\mathrm{tr}\,(D^3)\mathrm{tr}\,(D)\mathrm{tr}\,(C)^2\sum C_{ii_3}C_{i_3j} \\ =\ &\frac{N^2}{T^2}\tau(D^3)\tau(D)\tau(C)^2\sum C_{ii_3}C_{i_3j}. \end{aligned} \tag{3.9}$$

Consider the following identity

$$(C^{k_1+k_2})_{ij} = (C^{k_1})_{i1}(C^{k_2})_{1j} + \cdots + (C^{k_1})_{iN}(C^{k_2})_{Nj}. \tag{3.10}$$

We could write equation (3.9) as

$$q^2\tau(D^3)\tau(D)\tau(C)^2(C^2)_{ij}. \tag{3.11}$$

If N, T tends to infinity and we keep $N/T = q$ as a constant, the above limit would not vanish. Note that, in general, a solid loop will give us a $\tau(C^a)$, hence contribute a coefficient N and a dashed loop will give us a $\tau(D^b)$, hence contribute a coefficient T. Therefore, the limit of a single

diagram does not vanish if it has the same number of loops as the power of the $\frac{1}{T}$ in the front of the diagram.

The terms for $\langle E^3 \rangle_{ij}$ consists of three diagrams

The first two diagrams have enough loops, but the third one just have one loop. Hence the last one would vanish. Note that the last diagram has self-intersection, we call such diagram non-planar. We note that the diagram will vanish if and only if it is non-planar. Therefore, $\mathcal{G}_E$ equals to the infinite sum of all planar diagrams.

3.3 Dualization

Let us go back to the Figure 3.1, the outermost layer could be decomposed by equation (3.10). By rearranging the terms in equation (3.11) according to Figure.3.1, we obtain

$$\frac{N^2}{T^2}\left([\tau(D^3)\tau(C)^2C]\cdot[\tau(D)C]\right)_{ij}. \tag{3.12}$$

Note that

$$\begin{aligned}
\tau(C)^2\tau(D^3) &= \frac{\tau(C)^2}{T}\sum_{t_1}\left(\sum_{t_3}(D^2)_{t_1t_3}D_{t_3t_1}\right)\\
&= \frac{\tau(C)^2}{T}\mathrm{tr}\left(\sum_{t_3}(D^2)_{tt_3}D_{t_3s}\right)\\
&= \frac{1}{T}\mathrm{tr}\left(\sum_{t_3}\left([\tau(C)D]\cdot[\tau(C)D]\right)_{tt_3}D_{t_3s}\right)
\end{aligned} \tag{3.13}$$

Equation (3.13) looks like a mess, but its diagrammatic representation is straight forward.

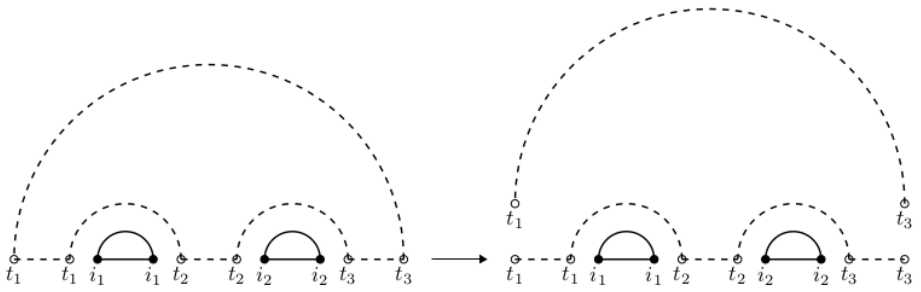


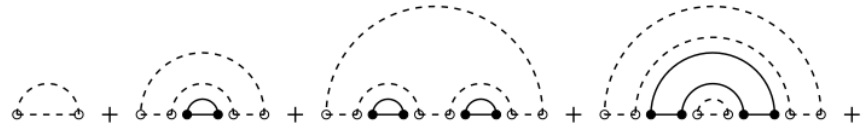
Figure 3.2: Operation $\mathrm{tr}\,(D\cdot)$

This decomposition inspires us to consider the diagrammatic dual of $\mathcal{G}_E$. We switch all the black circles, solid lines, and solid arcs with white circles, dashed lines, and dashed arcs respectively. Note that since for a single diagram, its number of solid lines and its number of dashed lines are not the same, we had better switch their positions in equation (2.18), that is

$$\mathcal{G}_{E*} = \left\langle \frac{I_T}{TI_T - z^{-1}X^TX} \right\rangle.$$

Hence, we find $\mathcal{G}_{E*}$ equals to

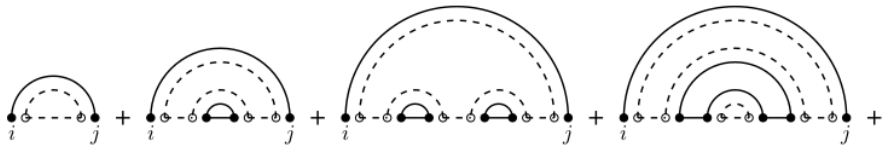
The argument above suggest us to consider $\mathrm{tr}\,(\mathcal{G}_{E*}D)$, and its diagrammatic representation is



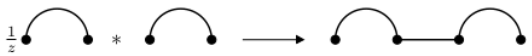
According to the definition given by the equation (3.2), we could give the diagram of

$$\Sigma_{ij} = \left(C\mathrm{tr}\,(D\mathcal{G}_{E*})\right)_{ij} = C_{ij}\mathrm{tr}\,(D\mathcal{G}_{E*}),$$

that is



Note that, the diagrams above do not have tail lines, this is because we want to get the right power of z and T. Σ is the series of all diagrams that its outermost layer only have one arc. For the diagram that has multiple arcs, we use the same methods we have used to deduce the equation (3.12). More specifically, by equation (3.10), matrix multiplication could be explained by the following diagram



Therefore, $\mathcal{G}_E$ is the geometric series respect to Σ. This proved the first part of equation (3.2). By applying a similar argument to the diagrammatic dual $\mathcal{G}_{E*}$ of $\mathcal{G}_E$, we could get that $\mathcal{G}_{E*}$ is the geometric series of the dual of Σ, which is $\Sigma_* = D\mathrm{tr}\,(C\mathcal{G}_E)$. This complete the proof of equations (3.2).

3.4 Recurrence Formulae for The Case $D = I_T$

By substitutions, equation (3.2) becomes equation (3.1), that is

$$\mathcal{G}_E(z) = \frac{I_N}{zI_N - C\mathrm{tr}\left(\dfrac{D}{I_T - D\mathrm{tr}\,(C\mathcal{G}_E(z))}\right)}.$$

If we set D to be the identity, then we have

$$\mathcal{G}(z)^{-1} = zI - \left(1 - T^{-1}\mathrm{tr}\,(C\mathcal{G}(z))\right)^{-1} C. \tag{3.14}$$

If we write

$$G_k = \langle E^k \rangle$$

for short, following equation (3.14), which rearranges to

$$(1 - q\tau(C\mathcal{G}(z))\,(z\mathcal{G}(z) - I) = C\mathcal{G}(z), \tag{3.15}$$

By comparing coefficients, we deduce a self-convolutive recurrence for the (G_k):

$$G_{k+1} = CG_k + q\sum_{j=1}^{k} \tau(CG_{j-1})G_{k+1-j}. \qquad k \geq 0 \tag{3.16}$$

Recall that in the end of last chapter, we had another recurrence formula, equation (2.21), that is

$$G_{k+1} = CG_k + q\sum_{j=1}^{k} \tau(G_j)CG_{k-j}, \qquad k \geq 0.$$

It is hard to see that they will generate the same moments, but they are actually the same. To show that (3.15) and (2.20) give the same answer, write $g = \tau(\mathcal{G}(z))$ and $\gamma = \tau(C\mathcal{G}(z))$ and take the trace of (3.15) to give

$$(1 - q\gamma)(zg - 1) = \gamma.$$

Hence

$$\gamma = (zg - 1)/(1 - q + qzg),$$

and so

$$(1 - q\gamma)(1 - q + qzg) = 1$$

which is what is needed to show that the two agree.

Chapter 4

Explicit Formulae

We had seen that $\langle E^k \rangle$ could be represented by the sum of planar diagrams. We are going to extend the diagrammatic method and obtain a more straightforward and quicker way to compute the moment $\tau(E^k)$. We are going to see that, by deforming the planar diagrams, we can transfer the problem of computing the moment into the problem of counting the non-crossing partitions. And each case in equation (1.2) correspond to a specific counting problem. And then we will use some technique from combinatorics and representation theory to solve these problems.

4.1 Non-Crossing Partition

If we take the trace of the diagrammatic representation of $\mathcal{G}_E$, there should be only loops in all diagrams. Again, we use the diagram in Figure 3.1 as the example. If we take the trace of equation (3.11), we simply get

$$q^2\tau(D^3)\tau(D)\tau(C)^2\tau(C^2).$$

To obtain its diagram, we can not just add an arc to Figure 3.1 since an extra arc will increase the power of C by 1. We should "glue" the two outer horizontal lines instead. But by doing this, the last dashed loop labeled by t_4 would be left alone. The connection of the outer lines suggest that we should consider a "circular" structure, which motivate the following construction.

We now contract all horizontal lines into distinct points. Then we place these points as the vertices of a k-gon in order. We say that two vertices are equivalent if there was an arc connect them. Since, we are considering the τ expression, arcs will form loops, thus this gives us an equivalence relation on the vertices. All the vertices in each equivalence class can give us a convex hull, we use solid (dashed) lines represent the edges of these convex hulls. We labeling the black vertices and white vertices respectively by $1, \ldots, k$. For example Figure 3.1 becomes

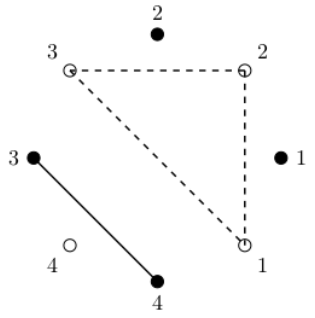


Figure 4.1: $\tau(D^3)\tau(D)\tau(C)^2\tau(C^2)$

To explain the diagram above, we need the definition of non-crossing partition.

Definition 4.1.1. *A partition π of a set $[k] = \{1, \ldots, k\}$ is a subset of its power set consisting of pairwise disjoint subsets whose union is the entire set. The elements in π are called blocks. We use $\mathrm{P}(k)$ to denote the set of all partitions on set $[k]$.*

Let $r = (r_1, \ldots, r_k)$ be a sequence satisfies $r_1 + 2r_2 + \cdots + kr_k = k$, we say a partition is of r-type if it has r_i blocks of size i.

A non-crossing partition σ is a partition of the vertices of a regular k-gon (labeled by the set $[k]$) with the property that the convex hulls of its blocks are pairwise disjoint. We use $\mathrm{NC}(k)$ to denote the set of all non-crossing partitions on set $[k]$.

Remark 4.1.2. *Suppose there are two distinct edges of a non-crossing partition, connect a_1 to a_2 and b_1 to b_2 respectively. Suppose $a_1 < a_2, b_1 < b_2$ and $a_1 < b_1$, then we must have $a_1 < b_1 < b_2 < a_2$ or $a_1 < a_2 < b_1 < b_2$.*

The equivalence classes on black vertices and white vertices give two non-crossing partitions on $[k]$, and these partitions has some relationship. If we have the solid partition, the dashed partition is uniquely determined, vice versa. This is called the Kreweras complement [14] and we will see a formal definition latter. Since the above construction is invertible, therefore, we have a one-to-one correspondence between all non-crossing diagrams appeared in $\langle E^k \rangle$ and $\mathrm{NC}(k)$, and hence, we could write $\tau(E)$ into sum of all possible non-crossing partitions.

For a given $\pi \in \mathrm{NC}(k)$ of r-type. Suppose we know the information of its Kreweras complement, say σ, is of s-type, such pair of partitions will contribute a term

$$q^{r_1+\cdots+r_k-1}\tau(C)^{r_1}\tau(C^2)^{r_2}\cdots\tau(C^k)^{r_k}\tau(D)^{s_1}\cdots\tau(D^k)^{s_k}. \tag{4.1}$$

Then the problem of computing the coefficient in the front of (4.1) becomes a counting problem.

4.2 Special Cases

Marčenko-Pastur case We first deal with the simplest case. The simplest case is when C and D are all identities. In this case, the counting problem has been simplified, $\tau(C^r)$ and $\tau(D^s)$ all equal to 1. Therefore, equation (4.1) become q^r, where r runs over all the integers from 0 to $k-1$. After combining the terms, the problem becomes to counting the partitions that has r blocks. This problem is well-known and has a very nice answer. The number of the non-crossing partition of set $[k]$ with r blocks is called Narayana number, and is given by

$$\mathrm{N}(k, r) = \frac{1}{k}\binom{k}{r}\binom{k}{r-1}. \tag{4.2}$$

Therefore, we have

$$\begin{aligned}
\tau(E^1) &= 1 \\
\tau(E^2) &= 1 + q \\
\tau(E^3) &= 1 + 3q + q^2 \\
&\vdots \\
\tau(E^6) &= 1 + 15q + 50q^2 + 50q^3 + 15q^4 + q^5.
\end{aligned}$$

This result is also mentioned in [13] chapter 3. In [15] chapter 2, they give the generating function of the Narayana number, which is

$$\sum_{k=1}^{\infty}\sum_{r=1}^{k}\mathrm{N}(k, r)z^k t^{r-1} = \frac{1 - z(t+1) - \sqrt{1 - 2z(t+1) + z^2(t+1)^2}}{2tz}. \tag{4.3}$$

By modifying this generating function, we will get exactly the same expression as equation (2.8).

Arbitrary C Now we discuss the case when C is arbitrary and $D = I_T$. Consider a non-crossing partition $\pi \in \mathrm{NC}(k)$, we define a sequence $(\lambda_1, \ldots, \lambda_k)_\pi$ associate to π by

$$\lambda_i = \begin{cases} |B| - 1 & \text{if } i = \min B \text{ for some block } B \in \pi; \\ -1 & \text{otherwise.} \end{cases}$$

We could see that this sequence have the following properties:

1. $\lambda_i \in \{-1, 0, 1, 2, \ldots\}$;
2. $\sum^{n} \lambda_i$ for all $n \leq k$;
3. $\sum^{k} \lambda_i = 0$.

Such sequence is called a Lukasiewicz path [14]. One can see that if we have a sequence $(\lambda_1, \ldots, \lambda_k)$ satisfies the properties above, we could recover the partition. The second property, positivity of the partial sum, guarantees the partition is non-crossing. Hence, there is a one-to-one correspondence between non-crossing partitions and Lukasiewicz paths.

Consider the following map:

$$\mathcal{R} \colon (\pi, \omega) \mapsto (\lambda_\omega, \ldots, \lambda_k, -1, \lambda_1, \ldots, \lambda_{\omega-1}),$$

where $\omega \in [k+1]$. This maps a non-crossing partition with a natural number to a sequence in $\{-1, 0, 1, 2, \ldots\}^{k+1}$ with its total sum equals to -1.

It has an inverse map. Consider a sequence

$$(\mu_1, \ldots, \mu_{k+1}) \in \{-1, 0, 1, 2, \ldots\}^{k+1} \quad \text{such that } \sum \mu_i = -1.$$

Let j be the first index such that $\sum_i^j \lambda_i$ is the least partial sum. We rearrange the sequence and get

$$(\mu_{j+1}, \ldots, \mu_{k+1}, \mu_1, \ldots, \mu_{j-1}) \in \{-1, 0, 1, 2, \ldots\}^{k}.$$

The total sum of our new sequence is 0, and all partial sum are all larger or equal to 0 by the choice of j, hence, this is a Lukasiewicz paths. This gives a inverse map of $\mathcal{R}$. Therefore, the image of $\mathcal{R}$ is

$$\left\{(\mu_1, \ldots, \mu_{k+1}) \in \{-1, 0, 1, 2, \ldots\}^{k+1} : \sum \mu_i = -1\right\},$$

and is a bijection.

This bijection allows us transfers our enumeration problem to the enumeration of sequence in $\{-1, 0, 1, 2, \ldots\}^{k+1}$ whose total sum equals to -1. More specifically, we require the sequence have i $r_i - 1$ for all i. This is a much easier thing to enumerate, and this equals to

$$\binom{k+1}{r_1, \ldots, r_k} = \frac{(k+1)!}{r_1! \cdots r_k! (k + 1 - \sum r_i)!} \tag{4.4}$$

We need to divide equation (4.4) by $k+1$, since $\mathcal{R}$ is a map from $\mathrm{NC}(k) \times [k+1]$. Hence, the coeficcient should be

$$\frac{1}{k+1}\binom{k+1}{r_1, \ldots, r_k} = \frac{k!}{r_1! \cdots r_k! (k + 1 - \sum r_i)!} \tag{4.5}$$

Then we could compute the moments and get

$$\begin{aligned} \tau(E^1) &= \tau(C) \\ \tau(E^2) &= \tau(C^2) + q\tau(C)^2 \\ \tau(E^3) &= \tau(C^3) + 3q\tau(C^2)\tau(C) + q^2\tau(C)^3 \\ \tau(E^4) &= \tau(C^4) + 2q\tau(C^2)^2 + 4q\tau(C^3)\tau(C) + 6q^2\tau(C^2)\tau(C)^2 + q^3\tau(C)^4, \end{aligned}$$

and the Stieltjes transform equals to

$$\mathfrak{g}_E(z) = \sum_{k=0}^{\infty} \sum_{\substack{r_1,\ldots,r_k \\ r_1+\cdots+kr_k=k}} \frac{1}{k+1}\binom{k+1}{r_1,\ldots,r_k}\tau(C)^{r_1}\tau(C^2)^{r_2}\cdots\tau(C^k)^{r_k}\frac{q^{r_1+\cdots+r_k-1}}{z^{k+1}}. \tag{4.6}$$

To solve the general case, we need to use some tools from lattice theory.

4.3 Lattice Theory and Kreweras Complement

Definition 4.3.1. *A partially ordered set (poset) $(L, \leq)$ consists of a set L with a relation $\leq$ has the following properties: for every $a, b, c \in L$*

1. *Reflexivity: $a \leq a$;*
2. *Antisymmetry: $a \leq b$ and $b \leq a$ then $a = b$;*
3. *Transitivity: $a \leq b$ and $b \leq c$ then $a \leq c$.*

We say a finite poset is a finite lattice, if it satisfies

1. *For any $a, b \in L$, the set $\{c \in L \colon a \leq c \text{ and } b \leq c\}$ is not empty and hence exist an minimal element we denoted by $a \vee b$ and call it the join of a, b;*
2. *For any $a, b \in L$, the set $\{c \in L \colon a \geq c \text{ and } b \geq c\}$ is not empty and hence exist an maximal element we denoted by $a \wedge b$ and call it the meet of a, b.*

A map $f \colon L \longrightarrow M$ between two lattice is called a lattice homomorphism if it satisfies for any $a, b \in L$

1. *$f(a \vee b) = f(a) \vee f(b)$;*
2. *$f(a \wedge b) = f(a) \wedge f(b)$.*

If it is a bijection then we call it a lattice isomorphism. If further $M = L$, then it is a lattice automorphism. We could see that all the automorphism of lattice L forms a group, and we use $\mathrm{Aut}(L)$ to denote it.

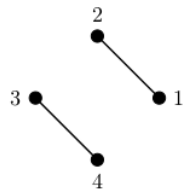
A map $\phi \colon L \longrightarrow M$ between two lattice is called a lattice anti-homomorphism if is satisfies for any $a, b \in L$

1. *$\phi(a \vee b) = \phi(a) \wedge \phi(b)$;*
2. *$\phi(a \wedge b) = \phi(a) \vee \phi(b)$.*

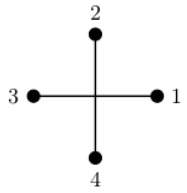
Similarly, we have anti-isomorphism and anti-automorphism. Note that the composition of two anti-homomorphisms gives us a lattice homomorohism and the composition of two anti-automorphisms gives us a lattice automorohism.

$\mathrm{NC}(k)$ has lattice structure. Set inclusion naturally define a partial order on $\mathrm{NC}(k)$. Intersection and union of two sets play the roles of join and meet of two elements. Hence $\mathrm{NC}(k)$ is a lattice. Similarly, we could see that $\mathrm{P}(k)$ is a lattice. However, $\mathrm{NC}(k)$ is not a sublattice of $\mathrm{P}(k)$, a lattice with the same meet and join operations.

Recall that symmetric group S_k has group action on k-gon. Therefore, each element of symmetric group give us a natural automorphism of $\mathrm{P}(k)$. However, S_k does not equal to $\mathrm{Aut}(\mathrm{NC}(k))$, we could see this by the following exmaple



acted by $(1\,4) \in S_4$ becomes



which is not in NC(4). Actually, we have

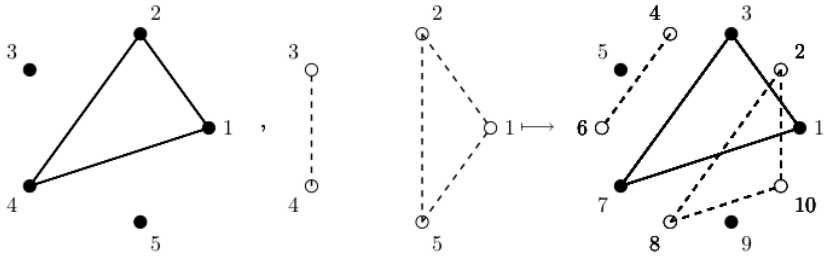
$$\mathrm{Aut}(\mathrm{NC}(k)) \cong D_{2k} = \langle \alpha, \beta \colon \alpha^k = \beta^2 = e, \beta\alpha\beta = \alpha^{-1} \rangle$$

the Dihedral group [14].

We need one more map to introduce the Kreweras complement. The following map is called the interlacing map

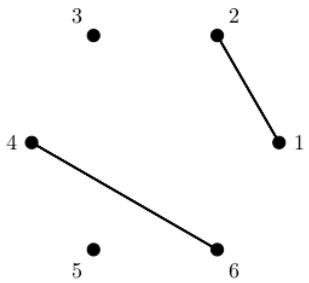
$$I \colon \mathrm{NC}(k) \times \mathrm{NC}(k) \longrightarrow \mathrm{P}(2k)$$

given by the following graph



Generally, the image of I is not in NC$(2k)$.

For a given $\pi \in \mathrm{NC}(k)$, we define $K_\pi = \{\sigma \in \mathrm{NC}(k) \colon I(\pi, \sigma) \in \mathrm{NC}(2k)\}$. For example, if π is given by



The lattice structure of K_π is showing in Figure 4.2.

K_π is a sublattice of NC(k). Since it is finite, it has a maximal element $\sup K_\pi$, we call it the Kreweras complement of π. We could naturally define a map on NC(k) to itself by

$$K \colon \mathrm{NC}(k) \longrightarrow \mathrm{NC}(k), \qquad \pi \mapsto \sup K_\pi.$$

K is an anti-automorphism. This could be seen from the fact that if $\pi \subseteq \sigma$, then $K(\pi) \supseteq K(\sigma)$, and then we have $K^2 = \alpha$. This leads to an important result

$$|\pi| + |K(\pi)| = k + 1. \tag{4.7}$$

Now consider an embedding

$$\mathrm{NC}(k) \hookrightarrow S_k$$

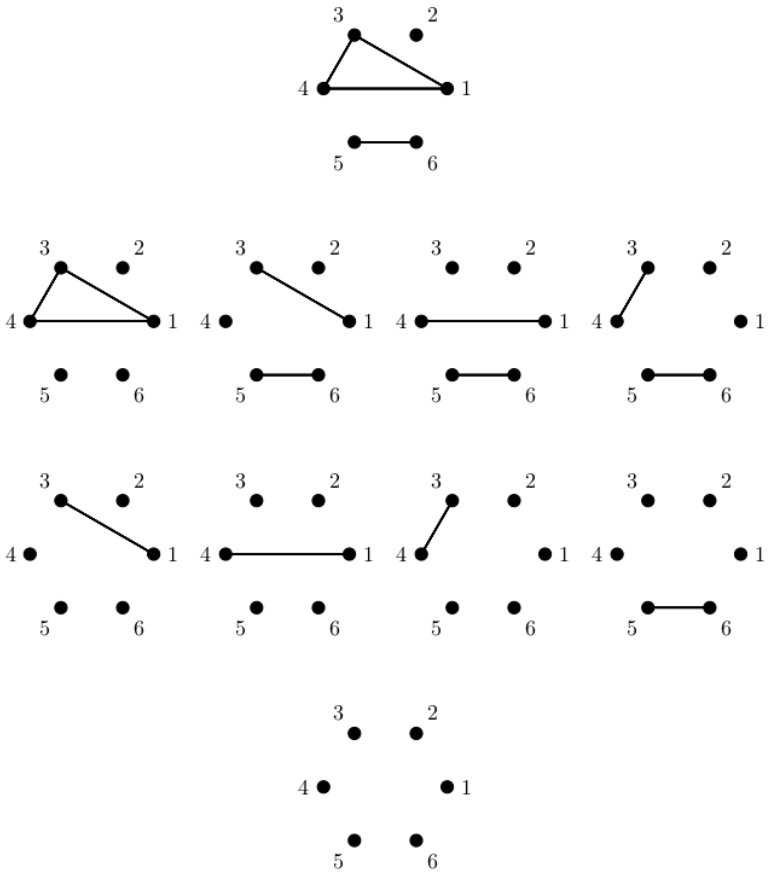


Figure 4.2: Lattice structure of K_π, the top one is the sup element, that is the Kreweras complement of π

defined as follow: for a block $B = \{a_1, \ldots, a_n\} \in \pi \mapsto (a_1\, a_2 \cdots a_n) \in S_k$ with $a_1 < a_2 < \cdots < a_n$, and π is mapped to the product of the image of its blocks. For convenience, when we say $\pi \in \mathrm{NC}(k)$, we default that π is an element of S_k. The anti-automorphism $K^2 = \alpha$ can be rewritten as $K^2(\pi) = \alpha \circ \pi \circ \alpha^{-1}$.

Another important result follows $K^2 = \alpha$ is

$$K(\pi) \circ \pi = \alpha = (1\, 2 \cdots k). \tag{4.8}$$

We could see this is true from the diagram. Suppose π maps a natural number a to another natural number b, hence, the line segment connect a and b is an edge of a block. In the diagram, the white vertex labeled by b is the one before (anti-clockwisely) the solid vertex labeled by b, it connected to the white vertex $a + 1$, this is due to the non-crossing property and the maximality of $K(\pi)$. Hence, $K(\pi)$ maps b to $a + 1$, which proves equation (4.8).

4.4 Representation Theory and Two Coloured Tree

Now we now that what is the Kreweras complement, but in order to enumerate it, we need to use some results from representation theory. These preparation will lead us to use theorem 2.2 of [5], which gives a formula to compute the connection coefficient. We will define later. We also want to argue that this coefficient is the answer to the counting problem for the general case: For

$r_1, \ldots, r_k, q_1, \ldots, q_k \in \mathbb{N}$, how many non-crossing partitions are there satisfy they have r_i blocks of size i and their Kreweras complement has s_i blocks of size i. [1]

Recall that we have mentioned the embedding of $\mathrm{NC}(k)$ into the symmetric group S_k. Consider the group ring $\mathbb{C}[S_k]$, recall that two elements $g, h \in S_k$ are conjugate if and only if they have the same circle type. We use $\mathrm{Cl}(g)$ to denote the conjugacy class of g. We define

$$K_g := \sum_{h \in \mathrm{Cl}(g)} h.$$

Then for some natural numbers $c^{\gamma}_{g_1 g_2}$, we have

$$K_{g_1} K_{g_2} = \sum_{\gamma} c^{\gamma}_{g_1 g_2} K_{\gamma}, \tag{4.9}$$

where γ runs over all conjugacy classes. We could write the left hand side into the linear sum since the center $Z(\mathbb{C}(S_k))$ is a group under multiplication and a vector space at the same time, the set $\{K_g\}$ forms a basis for $Z(\mathbb{C}(S_k))$. $c^{\alpha}_{K(\pi)\pi}$ is called the connection coefficient and we claim that it is the answer to the counting problem that we are looking for.[2] This claim is equivalent to say, for any $g, h \in S_k$ such that one of them is not in the image of $\mathrm{NC}(k)$ of the circle type $r_1, \ldots, r_k$ and $s_1, \ldots, s_k$ respectively, $hg \neq \alpha$.

Now we need a tool from combinatorics to decompose a cycle in S_k.

Definition 4.4.1. *A two-coloured edge-rooted tree T is a tree such that all of its edge has two different types of vertex, say white and black. One edge is thought as a rooted edge.*

A labeled two-coloured edge-rooted tree T is a two-coloured edge-rooted tree with all its edges are uniquely labeled by an element of set $[k]$, and the root edge is labeled by 1.

Suppose we have a two-coloured edge-rooted tree T and a cycle $(1\, a_2 \cdots a_k) \in S_k$, we give a "Depth-First" algorithm to traverse and label the two-coloured tree according to $(1\, a_2 \cdots a_k)$:

1. We start the transversal from the white vertex of the rooted edge;
2. Each time we travel from a white vertex to a black vertex, we label the edge by a natural number from $(1\, a_2 \cdots a_k)$ in order, hence, the rooted edge is labeled by 1;
3. When we reach the other end of an edge, we travel to the edge next to it anti-clockwise;
4. We stop this process if every edges are labeled.

Conversely, suppose we have a labeled two-coloured edge-rooted tree. For each black vertex, we write down the natural numbers labeled on the edge connect to the vertex in anti-clockwise order and get a cycle. By doing this to all black vertices, we get a product of cycle and hence get an element in S_k. Similarly, we can get another permutation from white vertices. For example, suppose we have the tree in Figure 4.3, we can get $\sigma = (1\,4)(6\,7\,9\,10)(11\,12\,13) \in S_k$ from white vertices and $\pi = (1\,5\,6\,11\,14)(2\,3\,4)(8\,9) \in S_k$ from black vertices. Note that $\sigma\pi = (1\,2 \cdots 14)$.

Recall that for a non-crossing partition π, we have $K(\pi)\pi = \alpha$. We claim that there is a one-to-one correspondence between the non-crossing partition and a two-coloured edge-rooted tree labeled by α.

We just need to show the permutation get from white vertices is non-crossing, the black vertices is non-crossing automatically follows the uniqueness of the Kreweras complement. Suppose we have a edge labeled by a_1 and the edge next to it anti-clockwise is labeled by a_2, thus we have $a_1 < a_2$. According to the algorithm, all the edges of the subtree connect to the black vertex of the edge a_1, would be labeled by the natural numbers between a_1 and a_2. Moreover, it contains all the natural numbers between a_1 and a_2. Note that the above argument is not true if α has been replaced by other permutation that is conjugate to α. Therefore, there is a one-to-one correspondence between the non-crossing partitions and the labeled two-coloured edge-rooted trees generated by α.

[1] If we know π, then $K(\pi)$ is uniquely determined, but just knowing $r_1, \ldots, r_k$ does not mean $q_1, \ldots, q_k$ is determined.

[2] π could be any partition satisfies the condition

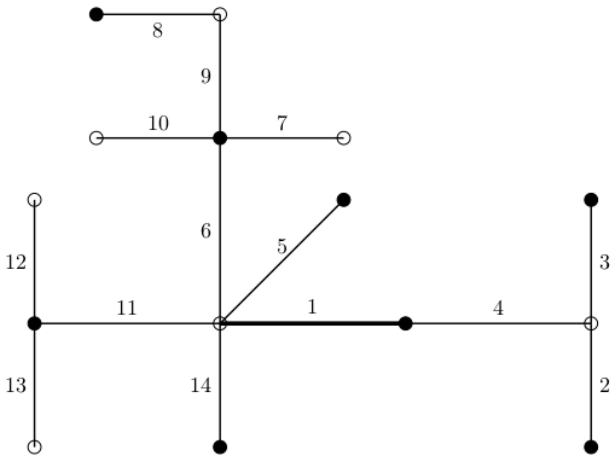


Figure 4.3: Example of a labeled two-coloured edge-rooted tree

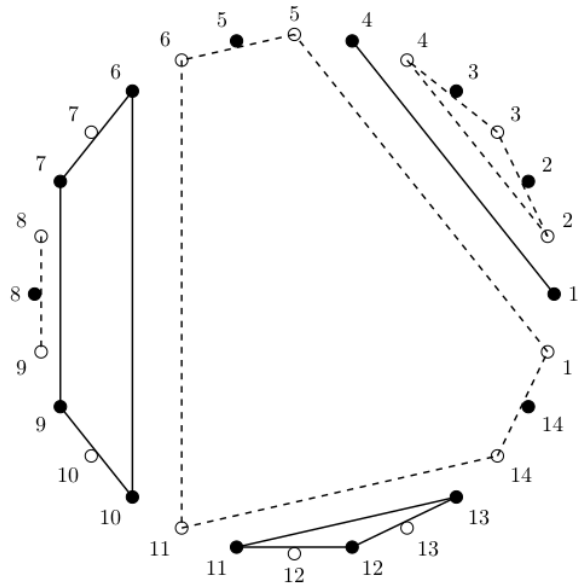


Figure 4.4: The tree in figure 4.3 corresponds to the pair of non-crossing partitions above

Remark 4.4.2. *To summarize, two-coloured plane edge-rooted trees with a permutation conjugate to α will give a labeled two-coloured plane edge-rooted trees, and then give a product decomposition of r,s-type(type given by tree). Two r,s-type non-crossing partitions those are mutually Kreweras complement will give a labeled two-coloured plane edge-rooted tree. This proves our claim on the connection coefficients.*

4.5 General Case

Now we are ready to apply the result from [5]. Theorem 2.2 in [5] states that if two permutation π, σ are of r-type and s-type respectively such that $s + r = k + 1$, then we have

$$c^{\alpha}_{\sigma\pi} = \frac{k}{(\sum r_i)(\sum s_i)} \binom{\sum r_i}{r_1, \ldots, r_k} \binom{\sum s_i}{s_1, \ldots, s_k} = k \frac{(\sum r_i - 1)!}{r_1! \ldots r_k!} \frac{(\sum s_i - 1)!}{s_1! \ldots s_k!}. \tag{4.10}$$

Then we could compute the moments and get

$$\begin{aligned}
\tau(E^1) &= \tau(C)\tau(D) \\
\tau(E^2) &= \tau(C^2)\tau(D)^2 + q\tau(C)^2\tau(D^2) \\
\tau(E^3) &= \tau(C^3)\tau(D)^3 + 3q\tau(C)\tau(C^2)\tau(D)\tau(D^2) + q^2\tau(C)^3\tau(D^3) \\
\tau(E^4) &= \tau(C^4)\tau(D)^4 + 2q\tau(C^2)^2\tau(D)^2\tau(D^2) + 4q\tau(C)\tau(C^3)\tau(D)^2\tau(D^2) \\
&\quad + 4q^2\tau(C)^2\tau(C^2)\tau(D)\tau(D^3) + 2q^2\tau(D^2)^2\tau(C)^2\tau(C^2) + q^3\tau(C)^4\tau(D^4)
\end{aligned}$$

Note that the coefficient of the moments are symmetric, this is because K is an anti-automorphism. We could write down the Stieltjes transform

$$\mathfrak{g}_E(z) = \sum_{k=0}^{\infty} \sum_{\substack{r_1, \ldots, s_k \\ r_1 + \cdots + k r_k = k \\ s_1 + \cdots + k s_k = k \\ r_1 + r_2 \cdots + s_{k-1} + s_k = k+1}} \frac{k q^{r_1 + \cdots + r_k - 1}}{(\sum r_i)(\sum s_i) z^{k+1}} \binom{\sum r_i}{r_1, \ldots, r_k} \binom{\sum s_i}{s_1, \ldots, s_k} \tau(C)^{r_1} \cdots \tau(D^k)^{s_k} \tag{4.11}$$

Chapter 5

Conclusion

The behaviour of the empirical covariance matrix $E = \frac{XX^T}{T}$ is determined by the true spatial covariance matrix C, the true temporal covariance matrix D and $q = N/T$, the ratio between the number of the variables and the number of the observations.

In chapter 2, we have discussed that if we know all the moments of a random matrix A, that is

$$\tau(A^k) = \frac{1}{N}\mathrm{tr}\,(A^k)$$

and if the Laurent series

$$\sum_{k=0}^{\infty} \frac{\tau(A^k)}{z^{k+1}}$$

converges around the infinity, We can then determine all of its eigenvalues uniquely. The method of computing the moments of E given by Bouchaud & Potters [1] involves computation of composition of Laurent series. Moreover, the computation of the moments for the empirical covariance matrix with non-trivial temporal structure involves the computation of the functional inverse of Laurent series, hence it is very difficult to compute though these methods.

In chapter 3, we have seen that according to the results of Sengupta & Mitra, we can give a new recurrence formula to compute the moments of E, and it is quicker. We also provided all missing details to the diagrammatic methods appear in [2].

In Chapter 4, we had extended the diagrammatic method and give some explicit formulae to compute the moments of the empirical covariance matrix E under different assumptions. For the general case, we use some results from representation theory given by Goulden & Jackson [5]. By using these formulas, we could compute the moments even more quickly. Moreover, for the case of $D = I$, the moment formula is invertible, that is, we could compute the moments of the true covariance matrix by the moments of the empirical matrix.

Now we would like to discuss some possible future extensions. Another transform in free probability theory, the R-transform $\mathcal{R}$, has additive structure. That is, for two random matrices A and B, we have $\mathcal{R}_{A+B} = \mathcal{R}_A + \mathcal{R}_B$. The coefficients of this R-transform is called free cumulants. According to some references, for example [14], the coefficients of the formula to compute the moments of a random matrix by free cumulants is the same as the coefficients of the moment formula for the case $D = I$. Note that it does not mean the $\tau(C^k)$ is the k-th cumulant of E, since the formula of the moments also contains q^n for some natural number n. But it may suggest some additive structure of covariance matrices.

Another possible extension is by considering the moment problem, that is how to compute the spectrum by the knowledge of the moments. We know how to compute the Laurent series expression of the Stieltjes transform near infinity, but the inverse formula (2.3) for the Stieltjes transform require the behaviour of the transform near the real line. Thus, we can not apply this formula directly. Any the numerical works on this direction may be very useful in practice.

Bibliography

[1] Marc Potters and Jean-Philippe Bouchaud. *A first course in random matrix theory for physicists, engineers and Data scientists*. Cambridge University Press, 2021.

[2] A. M. Sengupta and P. P. Mitra. Distributions of singular values for some random matrices. *Physical Review E*, 60(3):3389–3392, 1999.

[3] Z. Burda, A. Görlich, A. Jarosz, and J. Jurkiewicz. Signal and noise in correlation matrix. *Physica A: Statistical Mechanics and its Applications*, 343:295–310, 2004.

[4] Zdzisław Burda, Jerzy Jurkiewicz, and Bartłomiej Wacław. Spectral moments of correlated wishart matrices. *Physical Review E*, 71(2), 2005.

[5] I.P. Goulden and D.M. Jackson. The combinatorial relationship between trees, cacti and certain connection coefficients for the symmetric group. *European Journal of Combinatorics*, 13(5):357–365, 1992.

[6] Dan Voiculescu, Ken J. Dykema, and Andu Nica. *Free random variables: A noncommutative probability approach to free products with applications to random matrices, operator algebras and harmonic analysis on free groups*. American Math. Soc., 1992.

[7] Jean-Phillipe Bouchaud and Marc Potters. Financial applications of random matrix theory: A short review. *The Oxford Handbook of Random Matrix Theory*, page 823–850, 2015.

[8] Terence Tao. *Topics in random matrix theory*. American Mathematical Society, 2012.

[9] Hubert S. Wall. *Analytic theory of continued fractions*. Chelsea Publ. Co, 1973.

[10] Roland Speicher. Lecture notes on "random matrices", Sep 2020.

[11] V A Marčenko and L A Pastur. Distribution of eigenvalues for some sets of random matrices. *Mathematics of the USSR-Sbornik*, 1(4):457–483, 1967.

[12] Donald L. Cohn. *Measure theory second edition*. Springer New York, 2013.

[13] Zhidong Bai and Jack W. Silverstein. *Spectral analysis of large dimensional random matrices*. Science Press, 2006.

[14] Alexandru Nica and Roland Speicher. *Lectures on the combinatorics of free probability*. Cambridge University Press, 2006.

[15] T. Kyle Petersen. Eulerian numbers. *Birkhäuser Advanced Texts Basler Lehrbücher*, 2015.

[16] Richard P. Stanley. *Enumerative combinatorics*. Cambridge Univ. Press, 2008.