# Random Matrices: Theory and Practice

Yixiao Wang

April 2023

# Content

# What is Gaussian Ensembles?

- We now produce a $N \times N$ matrix H whose entries are independently sampled from a Gaussian probability density function with mean 0 and variance 1. To get real eigenvalues, the first thing to do is to symmetrize our matrix. Recall that a real symmetric matrix has N real eigenvalues.

## Definition

Symmetrizing $H_s = \frac{H+H^T}{2}$, then the eigenvalues of $H_s$ is real. We call $H_s$ is a sample of **Gaussian Orthogonal Ensembles(GOE)**.
Similarly, we can make the entries complex or quaternionic instead of real, then we get the **Gaussian Unitary Ensembles(GUE)** and **Gaussian Symplectic Ensembles(GSE)**, respectively.

# Layman's Classification

## Why Gaussian ensembles are important?

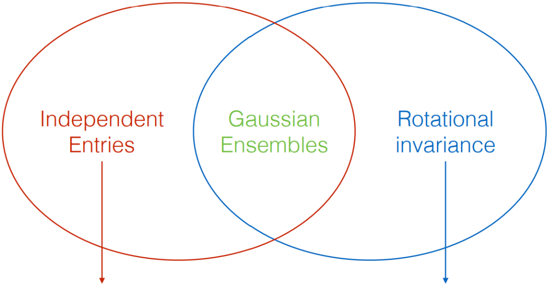
- The universe we are working on

$$X = \begin{cases} \text{real symmetric} \\ \text{complex hermitian} \\ \text{quaternion self-dual} \end{cases}$$

- **Independent entries:** the first group on the left gathers matrix models whose entries are independent random variables - modulo the symmetry requirements. Random matrices of this kind are usually called Wigner matrices
- **Rotational invariance:** the second group on the right is characterized by the so-called rotational invariance. In essence, this property means that any two matrices that are related via a similarity transformation $H' \stackrel{d}{=} UHU^{-1}$.

# Layman's Classification

## Why Gaussian ensembles are important?




$\rho[H] \propto \prod_{i=1}^{N} f_i(H_{ii}) \prod_{i<j} f_{ij}(H_{ij})$ $\qquad$ $\rho[H] = \rho[UHU^{-1}]$

Figure: Layman's classification

## Proposition(The orthogonal invariance of GOE)

Let $H_N$ be a random matrix drawn from GOE and $\forall U \in O(n)$, then the $UH_NU^{-1}$ is also a sample from GOE.

## Proof.

Since $H_N = \frac{H+H^T}{2}$, then we can get $UH_NU^{-1} = \frac{UHU^{-1}+UHU^{-1\,T}}{2}$. So we only need to prove

$$\begin{aligned} UHU^{-1} \sim H &\Leftrightarrow UH \sim H \\ &\Leftrightarrow UH_{*1} \sim H_{*1} \end{aligned}$$

,where the $H_{*1}$ is the first column of $H$. It is easy to see that the entries of $UH_{*1}$ are independent $\mathcal{N}(0,1)$ variables. □

# Wigner's Surmise

## Question

Consider a $2 \times 2$ GOE matrix $H_s = \begin{pmatrix} X_1 & X_3 \\ X_3 & X_2 \end{pmatrix}$, with $X_1, X_2 \sim \mathcal{N}(0,1)$ and $X_3 \sim \mathcal{N}(0,1/2)$. What is the pdf p(s) of the spacing $S = \lambda_2 - \lambda_1$ between its two eigenvalues $(\lambda_2 > \lambda_1)$?

## Proof.

The two eigenvalues are random variables, given in terms of the entries by the roots of the characteristic polynomial

$$\lambda^2 - Tr(H_s)\lambda + det(H_s)$$

therefore $\lambda_{1,2} = \frac{X_1 + X_2 \pm \sqrt{(X_1 - X_2)^2 + 4X_3^2}}{2}$ and $S = \sqrt{(X_1 - X_2)^2 + 4X_3^2}$.

## Wigner's Surmise

**Proof.**

It is easy to see that $X_1 - X_2, 2X_3$ $iid \sim \mathcal{N}(0,2)$, Changing variables as

$$\begin{cases} X_1 - X_2 = R\cos\Theta \\ 2X_3 = R\sin\Theta \end{cases}$$

and computing the corresponding Jacobian $J = \begin{vmatrix} \frac{\partial x_1 - x_2}{\partial r} & \frac{\partial x_1 - x_2}{\partial \theta} \\ \frac{\partial 2x_3}{\partial r} & \frac{\partial 2x_3}{\partial \theta} \end{vmatrix} = r$, one obtains

$$f_{R,\Theta}(r,\theta) = f_{(X_1 - X_2),2X_3}(r\cos\theta, r\sin\theta)\,|J| = \frac{r}{4\pi}e^{-\frac{r^2}{4}}$$

since S=R, thus

$$f_S(s) = \int_0^{2\pi} f_{S,\Theta}(s, theta) = \int_0^{2\pi} \frac{s}{4\pi}e^{-\frac{s^2}{4}} = \frac{s}{2}e^{-\frac{s^2}{4}} \square$$

# Wigner's Surmise

## Remark

We often rescale this pdf and define $\overline{f}(s) = \mathbb{E}(S) f(\mathbb{E}(S)s)$, such that $\int_0^\infty s\,\overline{f}(s) = 1$. For the GOE as above, $\overline{f}(s) = \frac{\pi s}{2}\exp(-\pi s^2/4)$, which is called "Wigner's surmise", whose plot is shown below.

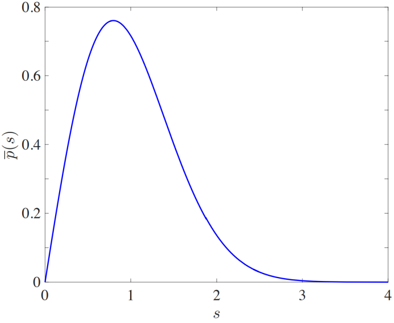



Figure: Wigner's surmise

# Wigner Hermitian matrix ensemble

## Definition

A Wigner Hermitian matrix ensemble is a random matrix ensemble $M_n = (m_{ij})_{1 \le i,j \le n}$ of Hermitian matrices (thus $m_{ij} = m_{ji}$; this includes real symmetric matrices as an important special case), in which the upper-triangular entries $m_{ij}, i > j$ are iid complex random variables with $\mathbb{E}(m_{ij}) = 0$ and $Var(m_{ij}) = 1$, and the diagonal entries $m_{ii}$ are iid real variables, independent of the upper-triangular entries, with $\mathbb{E}(m_{ii}) = 0$ and bounded variance. Also $\forall k \ge 2, \forall i, j$, the $k^{th}$ moment of $m_{ij}$ is existed.

## Particular special cases

Particular special cases of interest include the Gaussian Orthogonal Ensemble (GOE), the symmetric random sign matrices (aka symmetric Bernoulli ensemble), and the Gaussian Unitary Ensemble (GUE)

- We can generate T matrices from GOE(or GUE, GSE), collect the N (real) eigenvalues for each of them, and then produce a normalized histogram of the full set of $N \times T$ eigenvalues.We can get a plot for T = 50000 and N = 8 below.

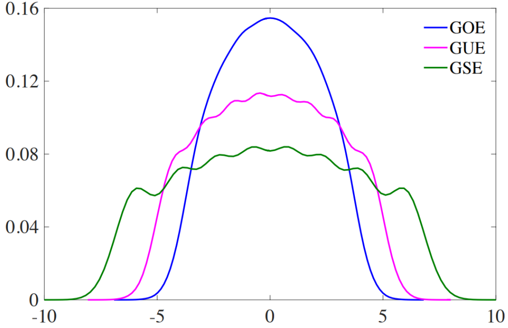
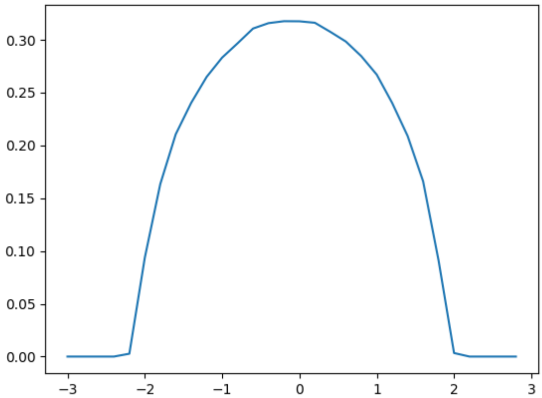



Figure 1.1: Histograms of GOE, GUE and GSE eigenvalues ($N = 8$ and $T = 50000$ samples).

# Wigner's Semicircle Law

- Now we can generate T $N \times N$ matrices whose upper triangle entries are independently and sampled from standard normal distribution, then produce a normalized histogram of the full set of $N \times T$ eigenvalues.We can get a plot for T = 500 and N = 200 below.



## Theorem

The jpdf of eigenvalues of a $N \times N$ Gaussian matrix is given by

$$f(x_1, \cdots, x_N) = \frac{1}{Z_{N,\beta}} e^{-\frac{1}{2}\sum_{i=1}^{N} x_i^2} \prod_{j<k} |x_j - x_k|^{\beta}$$

where

$$Z_{N,\beta} = (2\pi)^{N/2} \prod_{j=1}^{N} \frac{\Gamma(1 + j\beta/2)}{\Gamma(1 + \beta/2)}$$

is a normalization constant, enforcing $\int_{\mathbb{R}^N} f(x_1, \ldots, x_N) \prod_{j=1}^{N} \mathrm{d}x_j = 1$, and $\beta = 1, 2, 4$ is called the Dyson index, which is responding to the GOE, GUE and GSE respectively. Note that the eigenvalues are considered to be unordered here.

# Wigner's Semicircle Law

normalised matrix $\frac{1}{\sqrt{n}}M_n$

## Question

Given a $n \times n$ Hermitian matrix $M_n$, why we use $\frac{1}{\sqrt{n}}$ to normalise it?

## Solution

We prove that the eigenvalues of $M_n$ are typically of size $O(\sqrt{n})$. Let $X_n = \frac{M_n}{\sqrt{n}}$, consider that

$$\sum_{i=1}^{n} \lambda_i^2(X_n) = \operatorname{Tr} X_N^2 = \frac{1}{n} \operatorname{Tr} M_n^2 = \frac{1}{n} \sum_{i,\, j} m_{ij}^2$$

use the SLLN, we have $\frac{\sum_{i,\, j} m_{ij}^2 - n^2 \mathbb{E}(m_{ij}^2)}{n^2} \to 0$ ,thus $\bar{\lambda_i^2} \approx \mathbb{E}(m_{ij}^2) \sim 1$.

## Definition

Given any $n \times n$ Hermitian matrix $M_n$, we can form the (normalised) empirical spectral distribution (or ESD for short)

$$\mu_{M_n/\sqrt{n}} = \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i(M_n)/\sqrt{n}}$$

## Probability Measure

We can see that $\mu_{M_n/\sqrt{n}}$ is a probability measure by notice that $\mu_{M_n/\sqrt{n}}(\mathbb{R}) = 1$. Also for each set $A \in \mathcal{B}(\mathbb{R})$, $\mu_{M_n/\sqrt{n}}(A)$ is equal to the rate of eigenvalues in A (we only need to consider on $\mathbb{R}$ since the eigenvalues of a hermitian matrix are real).

# Wigner's Semicircle Law

## Wigner's Semicircle Law

Let $\sigma : \mathcal{B}(\mathbb{R}) \to [0,1]\,, A \mapsto \int_A \frac{\sqrt{4-x^2}}{2\pi}\mathrm{d}x$, then $\sigma \in Pr(\mathbb{R})$.
we want to prove that a sequence of random ESDs $\mu_{\frac{1}{\sqrt{n}}Mn}$ converge almost surely to $\sigma$.
i.e. $\forall$ test function $f \in C_c(\mathbb{R})$, the quantities $\int_{\mathbb{R}} f\,\mathrm{d}\mu_{M_n/\sqrt{n}}$ converge almost surely to $\int_{\mathbb{R}} f\,\mathrm{d}\sigma$.

## Analysis

Weierstrass's theorem tells us that any continuous function on a compact interval can be uniformly approximated by polynomials. Thus we only discuss about the quantities $\int_{\mathbb{R}} x^k\,\mathrm{d}\mu_{M_n/\sqrt{n}}$ converge almost surely to $\int_{\mathbb{R}} x^k\,\mathrm{d}\sigma$, $\forall k \in \mathbb{N}^*$.

# Wigner's Semicircle Law

## Analysis

Since

$$\int_{\mathbb{R}} x^k \,\mathrm{d}\mu_{M_n/\sqrt{n}} = \int_{\mathbb{R}} x^k \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i(X_n)}(\mathrm{d}x) = \frac{\sum_{i=1}^{n} \lambda_i^k}{n} = \frac{\operatorname{Tr} X_n^k}{n}$$

and we can derive that

$$\gamma_k := \int_{\mathbb{R}} x^k \,\mathrm{d}\sigma = \int_{-2}^{2} x^k \frac{\sqrt{4-x^2}}{2\pi} dx = \begin{cases} 0 & \text{, if k = 2m+1} \\ \frac{1}{m+1}\binom{2m}{m} & \text{, if k = 2m} \end{cases}$$

First we try to prove

$$\mathbb{E}\frac{\operatorname{Tr} X_n^k}{n} \to \gamma_k.$$

# Wigner's Semicircle Law

proof.

$$\mathbb{E}\frac{\operatorname{Tr} X_n^k}{n} = \frac{1}{n}\mathbb{E}\operatorname{Tr}\frac{M_n^k}{\sqrt{n}} = n^{-\left(\frac{k}{2}+1\right)}\mathbb{E}\operatorname{Tr}(M_n^k)$$

$$= n^{-\left(\frac{k}{2}+1\right)}\mathbb{E}\sum_{i=1}^{n}\sum_{\text{sequences of length k}} m_{ii_2}m_{i_2i_3}\ldots m_{i_ki}$$

$$= n^{-\left(\frac{k}{2}+1\right)}\mathbb{E}\sum_{\text{sequences of length k}} m_{i_1i_2}m_{i_2i_3}\ldots m_{i_ki_1}$$

Let's check each term in the sum.

## Wigner's Semicircle Law

**proof.**

We can view each term in the sum as a sequence whose length is 2m. Since we don't want the mean of a term equal to zero, we only need to consider a path where each edge gets traversed at least twice. Then there will be $\left[\frac{k}{2}\right]$ edges at most. We can prove that a connected graph with $\left[\frac{k}{2}\right]$ edges can only have no more than $\left[\frac{k}{2}\right]+1$ vertices. Furthermore, a connected graph must be a tree if it has $\left[\frac{k}{2}\right]$ edges and $\left[\frac{k}{2}\right]+1$ vertices.

# Wigner's Semicircle Law

## proof.

If $k$ is odd, we assume $k = 2m + 1$. Then there at most $m + 1$ vertices in the graph

$$\frac{1}{n^{m+\frac{3}{2}}} \binom{n}{m+1} f(m) \xrightarrow{n\to\infty} 0$$

, where $f(m)$ represents the mean of edges after fixed vertices.

If $k$ is even, we assume $k = 2m$.

If the vertices in the graph are less than $m + 1$, we assume there are $t\,(\leq m)$ different vertices in the graph, then the term will equal to

$$\frac{1}{n^{m+1}} \binom{n}{t} f(t) \sim O\left(\frac{1}{n^{m+1-t}}\right) \xrightarrow{n\to\infty} 0$$

, where $f(t)$ represents the mean of edges after fixed vertices.

# Wigner's Semicircle Law

## proof.

Now we only need to consider the graph that has $m$ edges and $m+1$ vertices exactly. For any ordered sequence (i.e. we have fixed the number of vertices also the visiting sequence). We reordered this sequence as $1, 2, \ldots, m+1$ we can let an

$$X : j \in \{1, 2 \ldots, 2m\} \mapsto \begin{cases} 1 & \text{, if jth step vist a new vertex} \\ -1 & \text{, otherwise} \end{cases}$$

let $S(j) = \sum_{i=1}^{j} X(i)$, then we get a simple random walk in 1 dimension. We can easy to find that $S(j) \geq 0, \forall j$. Furthermore, we can make a bijective map called $\mathcal{F}$ from all these types of paths to $(X(1), \ldots, X(2m))$ by definition.

# Wigner's Semicircle Law

**proof.**

We prove $\mathcal{F}$ is bijective first.
To begin with, we prove that we can form a path from any given sequence satisfy $X(i) = \pm 1$, $S(i) \geq 0$ with a length of 2m. Let $T = \{1, 2 \ldots, m+1\}$, $T_1 = T$ and $A$ will be a sequence with length 2m to record the vertices traversed by. For all $i \in \{1, 2 \ldots, 2m\}$, if $X(i) = 1$, let $A(i) = \min T_i$, $T_{i+1} = T_i \setminus \{A(i)\}$.Otherwise we find $j$ satisfy that $A(j) = A(i-1)$ and $X(j) = 1$, then let $A(i) = A(j-1)$. We can get a path from the sequence $A$ ($A(i)$ represents the vertex after $i$ steps).
On the one hand, for each sequence in $(X(1), \ldots, X(2m))$ we can find a path to be mapped to, thus $\mathcal{F}$ is surjective.

# Wigner's Semicircle Law

**proof.**

On the other hand, if we have two paths named $A, B$ ($A(i)$ and $B(i)$ represents the vertex after $i$ steps), but $\mathcal{F}(A) = \mathcal{F}(B)$. We assume $i$ is the first index such that $A(i) \neq B(i)$. By definition, $X(i) \neq 1$, since the order of the sequence is fixed, thus $X(i) = -1$. However the graph is a tree, there will be only one path from the beginning to $A(i-1) = B(i-1)$, then the edge connects from $A(i-1) = B(i-1)$ and only traversed once is determined. Other edges connected from $A(i-1)$ have been traversed twice or never. So we can derive that $A(i) = B(i)$, thus $\mathcal{F}$ is injective. As a result, $\mathcal{F}$ is bijective.

### proof.

Now the problem is change to a problem of random walk.

$$\begin{aligned}\#\{S(i) \geq 0 \forall i \in \{1,2,\dots,2m\}, S(0)=S(2m)=0\} &= \binom{2m-2}{m-1} - \binom{2m-2}{m-3} \\ &= \frac{1}{m+1}\binom{2m}{m}\end{aligned}$$

Thus

$$\begin{aligned}&n^{-\left(\frac{k}{2}+1\right)}\mathbb{E}\sum_{\text{sequence of length k}} m_{i_1 i_2} m_{i_2 i_3} \dots m_{i_k i_1} \\ =&\frac{1}{n^{m+1}}\binom{n}{m+1}(m+1)!\frac{1}{m+1}\binom{2m}{m} + O(\frac{1}{n}) \\ \rightarrow&\frac{1}{m+1}\binom{2m}{m} \quad \text{as } n \rightarrow \infty\end{aligned}$$

**proof.**

Next we consider the differ of $\frac{\operatorname{Tr} X_n^k}{n}$ and $\mathbb{E}\frac{\operatorname{Tr} X_n^k}{n}$. We first consider $Var\frac{\operatorname{Tr} X_n^k}{n}$

$$Var\frac{\operatorname{Tr} X_n^k}{n} = \frac{1}{n^{k+2}} \sum \mathbb{E} Y_i Y_j - \mathbb{E} Y_i \mathbb{E} Y_j$$

where $Y_i$ represents a sequence with length k. we named the set of vertex of $Y_i, Y_j$ as $I$, $J$. Now we consider how many paths such that $\mathbb{E} Y_i Y_j - \mathbb{E} Y_i \mathbb{E} Y_j \neq 0$.

# Wigner's Semicircle Law

## proof.

Now we consider how many different vertices in $I \cup J$. Since $Y_i$, $Y_j$ include 2k steps, if there is an edge has been traversed only once, then $\mathbb{E}Y_iY_j = \mathbb{E}Y_i\mathbb{E}Y_j = 0$. Thus each edge has been traversed at least twice. As a result, there are k edges at most.

If $\#\{I \cup J\} \geq k+2$, then the graph is not connected, but $Y_i$, $Y_j$ are connected respectively, so they are two independent paths. Then $\mathbb{E}Y_iY_j - \mathbb{E}Y_i\mathbb{E}Y_j = 0$.

If $\#\{I \cup J\} = k+1$, then the graph is a tree. We assume that if there exists an edge traversed by both $Y_i$, $Y_j$. As a result, this edge only traversed in $Y_i$ once, however, $Y_i$ is also a tree(since the connected subgraph of a tree is also a tree), which means it has only one way from the beginning to that edge and can get back to the beginning, which means that edge must be traversed twice! Thus $Y_i$, $Y_j$ are two independent paths as well.

# Wigner's Semicircle Law

**proof.**

So we can only consider $\#\{I \cup J\} \leq k$, and at this time

$$
\begin{aligned}
Var\frac{\operatorname{Tr} X_n^k}{n} &= \frac{1}{n^{k+2}} \sum \mathbb{E} Y_i Y_j - \mathbb{E} Y_i \mathbb{E} Y_j \\
&\sim \frac{1}{n^{k+2}} \binom{n}{k} f(k) \\
&\sim O(\frac{1}{n^2})
\end{aligned}
$$

where $f(k)$ represents the sum of each term after fixed vertices.

# Wigner's Semicircle Law

**proof.**

Finally, we prove that $\frac{\operatorname{Tr} X_N^k}{n} \xrightarrow{a.s.} \gamma_k$.By B-C lemma we have

$$
\begin{aligned}
&\frac{\operatorname{Tr} X_N^k}{n} \xrightarrow{a.s.} \gamma_k \\
\Leftrightarrow &\mathbb{P}\left(\left|\frac{\operatorname{Tr} X_N^k}{n} - \gamma_k\right| > \epsilon \quad i.o.\right) = 0 \quad \forall \epsilon > 0 \\
\Leftarrow &\sum \mathbb{P}\left(\left|\frac{\operatorname{Tr} X_N^k}{n} - \gamma_k\right| > \epsilon\right) < \infty
\end{aligned}
$$

# Wigner's Semicircle Law

**proof.**

Infact, by Markov Inequality

$$
\begin{aligned}
&\sum \mathbb{P}\left(\left|\frac{\operatorname{Tr} X_N^k}{n} - \gamma_k\right| > \epsilon\right) \\
&\leq \sum \mathbb{P}\left(\left|\frac{\operatorname{Tr} X_N^k}{n} - \mathbb{E}\frac{\operatorname{Tr} X_N^k}{n}\right| > \epsilon/2\right) + \mathbb{P}\left(\left|\mathbb{E}\frac{\operatorname{Tr} X_N^k}{n} - \gamma_k\right| > \epsilon/2\right) \\
&\leq \frac{4}{\epsilon^2}\sum Var(\frac{\operatorname{Tr} X_N^k}{n}) + C \\
&\sim \frac{4}{\epsilon^2}\sum \frac{1}{n^2} + C \\
&< \infty
\end{aligned}
$$

$\square$

## Introduction

Let $(p(n))_{n\geq 1}$ be a sequence of positive integers such that $\lim_{n\to\infty} \frac{p(n)}{n} = \alpha \geq 1$. Consider the $np(n)$ matrix $X_n$ whose entries are i.i.d. of mean 0 and variance 1, and with the kth moment bounded by some $r_k < \infty$ not depending on n.

As before, we will actually study the normalized matrix $Yn := \frac{X_N}{\sqrt{n}}$. The Marcenko-Pastur law is concerned with the distribution of the singular values of $Y_n$, which by definition are the eigenvalues of the $n \times n$ Wishart matrix $Wn = Y_n Y_n^T \in \mathbb{R}^{n\times n}$ .

As with the semicircle law, the limiting behaviour of these eigenvalues can be understood by considering the empirical spectral distribution $\mu_n$ of a Wishart matrix $W_n$ as $n \to \infty$.

## Definition

Given any $n \times n$ Wishart matrix $W_n$, we can form the empirical spectral distribution

$$\mu_{W_n} = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i}$$

where $\lambda_i$ are the eigenvalues of the Wishart matrix.

# Marcenko-Pastur Law

## Marcenko-Pastur Law

Let $\sigma : \mathcal{B}(\mathbb{R}) \to [0,1]\,, A \mapsto \int_A \frac{\sqrt{(\lambda_+ - x)(x-\lambda_-)}}{2\pi x} \mathbb{1}_{\lambda_- \le x \le \lambda_+} \mathrm{d}x$, where $\lambda_+ = (1+\sqrt{\alpha})^2, \lambda_- = (1-\sqrt{\alpha})^2$, then $\sigma \in Pr(\mathbb{R})$.
we want to prove that a sequence of random ESDs $\mu_{W_n}$ converge almost surely to $\sigma$.

## proof.

In oder to discuss the method of moment we only prove

$$\frac{1}{n}\mathbb{E}\,\mathrm{Tr}\, W_n^k \to \int x^k \mathrm{d}\sigma$$

# Marcenko-Pastur Law

## proof.

$$\begin{aligned}
\frac{1}{n}\mathbb{E}\operatorname{Tr} W_n^k &= \frac{1}{n^{k+1}}\mathbb{E}\operatorname{Tr}\left(X_n X_n^T\right)^k \\
&= \frac{1}{n^{k+1}} \sum_{\text{sequence of length k}} \mathbb{E}\left(X_n X_n^T\right)_{i_1 i_2} \ldots \left(X_n X_n^T\right)_{i_k i_1} \\
&= \frac{1}{n^{k+1}} \sum_{i_1 \ldots i_k; j_i \ldots j_k} \mathbb{E} X_{i_1 j_1} X_{i_2 j_1} X_{i_2 j_2} X_{i_3 j_2} \ldots X_{i_k j_k} X_{i_1 j_k}
\end{aligned}$$

As before, we focus on $\mathbb{E} X_{i_1 j_1} X_{i_2 j_1} X_{i_2 j_2} X_{i_3 j_2} \ldots X_{i_k j_k} X_{i_1 j_k}$. We can think of each such term as a connected bipartite graph on the sets of vertices $\{i_1, \ldots, i_k\}$ and $\{j_1, \ldots, j_k\}$, where the total number of edges (with repetitions) is $2k$.

# Marcenko-Pastur Law

## proof.

Suppose $n_i$ and $n_j$ denote the number of distinct $i$ vertices and $j$ vertices, respectively. Since each edge needs to be traversed at least twice, there are at most $k+1$ distinct vertices, so $n_i + n_j \leq k+1$.
We notice that if $n_i + n_j < k+1$, such terms will equal to

$$\frac{1}{n^{k+1}} \binom{n}{n_i + n_j} f(k) \xrightarrow{n \to \infty} 0$$

where $f(k)$ represents different kinds of bipartite graphs after fixed vertices. If $n_i + n_j = k+1$ there are $k$ unique edges and the resulting graph is a tree. Such terms will become the dominant ones in the sum in the limit $n \to \infty$.

# Marcenko-Pastur Law

## proof.

Now we focus on $n_i + n_j = k + 1$.
Notice that there are $n(n-1)\dots(n-n_i+1)p(p-1)\dots(p-n_j+1)$ corresponding choices for $i$ and $j$. Because $p \approx n\alpha$ for large n and we are in the case $n_i + n_j = k + 1$, the number of choices is asymptotically equal to $n^{k+1}\alpha^{n_j}$.
Let $\beta_k$ represents the number of type sequence of length 2k with weight after dividing by $n^{k+1}$.

$$\beta_k = \sum_{\text{type sequence of length 2k}} \alpha^{n_j}$$

Thus we can decuce

$$\frac{1}{n}\mathbb{E}\,\mathrm{Tr}\,W_n^k = \beta_k$$

## proof.

The goal is to establish a recurrence relation between the $\beta_k$ in order to compute the general term.

We assume that we first go back to the beginning after the $2j$ steps. Then each edge that is traversed in first $2j$ steps will be traversed twice, since this bipartite graph is also a tree. Thus if we delete the first step and the 2jth step, we will have a sequence with length $2(j-1)$. This sequence represents the number of type sequences of length $2(j-1)$ with weight, which begin from $\{1, 2 \ldots, p\}$. We named it $\gamma_{j-1}$. Then we have

$$\beta_k = \sum_{j=1}^{k} \gamma_{j-1} \beta_{k-j}$$

# Marcenko-Pastur Law

**proof.**

Similarly

$$\gamma_k = \sum_{j=1}^{k} \beta_{j-1}\gamma_{k-j}$$

Also we have $\gamma_0 = \alpha, \beta_0 = 1(\gamma_0, \beta_0$represents we only chose one points in $\{1, \ldots, p\}, \{1, \ldots, n\}$ respectively without walking)
Then we can deduce that

$$\beta_k = \gamma_k$$

Thus

$$\beta_k = (\alpha - 1)\beta_{k-1} + \sum_{j=1}^{k} \beta_{j-1}\beta_{k-j}$$

# Marcenko-Pastur Law

## proof.

In particular, if $\hat{\beta}(x) := \sum_{k=0}^{\infty} \beta_k x^k$ is the generating function for the $\beta_k$, the previous identity leads to the following equality for $\hat{\beta}$:

$$\hat{\beta}(x) = 1 + x\hat{\beta}(x)^2 + (\alpha - 1)x\hat{\beta}(x)$$

For convenience, we only see what happens when $\alpha = 1$, by Taylor-Expansion

$$\begin{aligned}
&\hat{\beta}(x) = 1 + x\hat{\beta}(x)^2 \\
\Rightarrow &\hat{\beta}(x) = \frac{1 - \sqrt{1 - 4x}}{2x} \\
\Rightarrow &\beta_k = \frac{1}{2}(-1)\frac{\binom{\frac{1}{2}}{k+1}}{(k+1)!}(-4)^{k+1} = \frac{1}{2k+1}\binom{2k}{k}
\end{aligned}$$

□

### Yau 2021 Problem 4

Let $X^n = X_{ij}$ be an $n \times n$ random matrix whose entries are independent and identically distributed random variables with the symmetric Bernoulli distribution $\mathbb{P}\{X = 0\} = \mathbb{P}\{X = 1\} = \frac{1}{2}$. Let $p_n = \mathbb{P}\{\det(X^n) \text{ is odd}\}$. Show that $\lim_{n\to\infty} p_n > 0$.

# Examples and Solutions

## Yau 2016 Problem 2

Let $X$ be a $N \times N$ random matrix with *i.i.d.* random entries, and $\mathbb{P}(X_{11} = 1) = \mathbb{P}(X_{11} = -1) = 1/2$ Define

$$\|X\|_{op} = \sup_{v \in \mathbb{C}^N : \|\vec{v}\|_2 = 1} \|X\vec{v}\|_2$$

Please show that for any fixed $\delta > 0$,

$$\lim_{N \to \infty} \mathbb{P}(\|X\|_{op} \geq N^{1/2+\delta}) = 0$$

Hint: $\|X\|_{op}^2 \leq \operatorname{Tr}|X|^2$

# The End