**Introduction to Random Matrix Theory and Applications**

**Def:** Let \( \{Z_{ij}\}_{i,j \in \mathbb{N}} \) be independent random variables, such that:

- \( \mathbb{E}[Z_{ij}] = 0 \) for all \( i \neq j \).
- \( \mathbb{E}[Z_{ij}^2] = 1 \) for all \( i \neq j \). (No requirement for \( i = j \)).
- \( \mathbb{E}[|Z_{ij}|^{4k}] < C_k \) for a constant \( C_k \) and all \( i \neq j \).

The symmetric matrix \( X \in \mathbb{R}^{n \times n} \) with

\[
X_{ij} = \begin{cases} 
  Z_{ij} / \sqrt{n} & \text{if } i < j \\
  0 & \text{if } i = j \\
  Z_{ji} / \sqrt{n} & \text{if } i > j
\end{cases}
\]

is called a Wigner matrix.

**Def:** If \( Z_{ij} \sim \mathcal{N}(0,1) \) for \( i < j \) and \( Z_{ij} \sim \mathcal{N}(0,2) \) for \( i = j \),

then \( X \) is called the Gaussian Orthogonal Ensemble (GOE).

We write \( X \sim \text{GOE}(n) \).

**Prop:** If \( X \sim \text{GOE}(n) \) and \( O \in \mathbb{R}^{n \times n} \) is any orthogonal matrix,

then also \( O^* X O \sim \text{GOE}(n) \).

**Proof:** Let \( W \in \mathbb{R}^{n \times n} \) be an asymmetric matrix,

with \( W_{ij} \sim \mathcal{N}(0,1) \) for all \( i,j \in \{1, \ldots, n\} \). Then \( X = \frac{W}{\sqrt{n}} \) is a symmetric matrix with Gaussian entries, satisfying:

- \( \forall X_{ij} : i \neq j \) are independent,
- \( \mathbb{E}[X_{ij}] = 0 \) for all \( i \neq j \),
- \( \mathbb{E}[X_{ij}^2] = \frac{1}{2n} \mathbb{E}[(W_{ij} + W_{ji})^2] = 0 \).
Thus \( w \sim \text{GOE}(n) \). Let \( w = w_0w_0^T \). Then \( O^T X O = w_0 w_0^T \), and it suffices to show \( w \sim \hat{w} \). For this, note:

\[
\frac{1}{n} \sum w_{ij}^2 = \text{Tr} \frac{1}{n} w w^T = \text{Tr} \hat{w} \hat{w} = \frac{1}{n} \sum \hat{w}_{ij}^2.
\]

So \( w \sim O^T w_0 \) is an isometry from \( \mathbb{R}^n \) to itself, with Jacobian 1.

The joint density of entries of \( w \) is

\[
\frac{1}{(2\pi)^{n^2/2}} e^{-\frac{1}{2} \frac{1}{n} \sum w_{ij}^2} = \frac{1}{(2\pi)^{n^2/2}} e^{-\frac{1}{2} \frac{1}{n} \sum \hat{w}_{ij}^2}.
\]

Thus \( \frac{1}{(2\pi)^{n^2/2}} e^{-\frac{1}{2} \frac{1}{n} \sum \hat{w}_{ij}^2} \) is the joint density of entries of \( \hat{w} \), i.e., \( \hat{w} \sim \text{GOE}(n) \).

Corollary: Let \( X \sim \text{GOE}(n) \). Let \( X = V \Lambda V^T \) be its spectral decomposition, where \( \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n) \) is ordered s.t. \( \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \). Then for any orthogonal matrix \( O \in \mathbb{R}^{n \times n} \), \( O^T X O \) is also \( \text{GOE}(n) \).

**Def:** For \( X \in \mathbb{R}^{n \times n} \) symmetric, its empirical spectral distribution (ESD) is the discrete probability distribution

\[
\mu_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i}(x),
\]

where \( \lambda_1, \ldots, \lambda_n \) are the (real) eigenvalues of \( X \).

Thus, let \( X \in \mathbb{R}^{n \times n} \) be Wigner. Then almost surely as \( n \to \infty \), its ESD converges weakly to the semicircle law, as in [E22]. This is the continuous distribution with density

\[
F(x) = \frac{1}{2} \sqrt{4 - x^2}; \quad \text{if} \quad x \in [-2, 2].
\]

Probability measure: weak convergence, and the moment method.

Idea: We want a way of describing when two probability measures in and \( \mu \) are “close”. We often do this by showing that

\[
| \int f(x) d\mu(x) - \int f(x) d\rho(x) | \text{ is small}
\]

for a sufficiently large class of test functions \( f \).

In random matrix theory, we want to study a class of test functions \( f \) such that

- Sufficiently large to approximate any other “nice” function.

\[
| \int f(x) d\mu(x) | = \frac{1}{n} \sum_{i=1}^{n} f(\lambda_i) \text{ can be described simply}
\]

in terms of the entries of \( X \).

Today: We'll take polynomials, generated by \( E(x) = x^k \).

Then \( \int f(x) d\mu(x) = \frac{1}{n} \sum X_i x^k = \frac{1}{n} \text{Tr} X^k \).

This is called the moment method.

Next week: We'll take rational functions, generated by \( E(x) = \frac{1}{x} \).

For \( \phi \in \mathcal{C} \), Then \( \int \phi(x) d\mu(x) = \frac{1}{n} \sum_{i=1}^{n} \phi(\lambda_i) = \frac{1}{n} \text{Tr} \phi(X) \).

This is called the Stripes transform/Resolvent method.
**Def:** Let $P$ be the space of (Borel) probability measures on $\mathbb{R}$.

$\mu_n \in P$ converges weakly to $\mu \in P$ if for any continuous, bounded function $f: \mathbb{R} \to \mathbb{R},$

$$\int f(x) \, d\mu_n(x) \to \int f(x) \, d\mu(x) \text{ as } n \to \infty.$$  

**Facts:** This topology of weak convergence is metrizable. There is a unique $d(\cdot, \cdot)$ on $P$ such that $\mu_n \to \mu$ weakly if and only if $d(\mu_n, \mu) \to 0$.

One such choice is the Levy distance

$$d(\mu_n, \mu) = \inf \{ \varepsilon : F_n(x-x) - \varepsilon \leq F(x) \leq F_n(x+x) + \varepsilon \},$$  

where $F_n, F$ are the CDFs of $\mu_n, \mu$.

$\mu_n \to \mu$ weakly is equivalent to $F_n(x) \to F(x)$ at every $x$ where $F$ is continuous.

The open sets of $(P, d)$ generate a Borel $\sigma$-algebra on $P$.

For a probability space $(X, \mathcal{F}, P)$, a random measure $\mu$ is a Borel-measurable map $X \to P$. We'll typically ignore $X$ and just write $\mu \in P$ (in the same way we write $X \in \mathcal{F}$ for a fixed $X$).

Then $\mu_n \to \mu$ weakly $\iff d(\mu_n, \mu) \to 0$ as $n \to \infty$.

**Prop (Almost-Matching):** Suppose $\int x^k \, d\mu_n(x) \to \int x^k \, d\mu(x)$ almost surely for each fixed integer $k \geq 0$, and $\mu \in P$ has bounded support. Then $\mu_n \to \mu$ weakly.

**Proof:** Let the support of $\mu$ be contained in $[-M, M]$. Then for any $i$ fixed

$$\int x^i \, d\mu_n(x) \leq \int x^i \, d\mu(x) \leq (M+1)^i.$$

Then

$$\frac{\int x^i \, d\mu_n(x)}{(M+1)^i} \leq \left( \frac{M}{M+1} \right)^i.$$

Since $k$ is arbitrary, this shows $\int x^k \, d\mu_n(x) \to 0$ for any fixed $k$.

Now let $f(x)$ be any continuous, bounded function.

Fix $\varepsilon > 0$. By the Weierstrass approximation theorem, there is a polynomial $p(x)$ s.t.

$$\sup_{|x| \leq M+1} |p(x) - f(x)| < \varepsilon.$$

Then

$$\int p(x) \, d\mu_n(x) - \int p(x) \, d\mu(x)$$

$$\leq \int p(x) \, d\mu_n(x) - \int p(x) \, d\mu_n(x)$$

$$+ \int |p(x) - f(x)| \, d\mu_n(x) + \int |p(x) - f(x)| \, d\mu_n(x).$$

1. $\int p(x) \, d\mu_n(x) \to \int p(x) \, d\mu_n(x) \to 0.$
2. $\int |p(x) - f(x)| \, d\mu_n(x) = \int |p(x) - f(x)| \, d\mu_n(x) \leq \varepsilon.$
3. $\int |p(x) - f(x)| \, d\mu_n(x) \leq \int |p(x) - f(x)| \, d\mu_n(x) \leq \varepsilon.$

Since $\varepsilon > 0$ is arbitrary, this shows $\int f(x) \, d\mu_n(x) \to \int f(x) \, d\mu(x)$.

**Note:** Some condition on $\mu$ is necessary, otherwise it may not be uniquely defined by its moments.

**Cor:** If $\mu \in P$ is random, $\int x^k \, d\mu(x) \to \int x^k \, d\mu(x)$ a.s. almost surely and $\mu$ has...
MOMENTS OF THE SEMICIRCLE LAW

Q: Let \( p_k \) be the semicircle law on \([-2,2]\). What is \( m_k = \int_{-2}^{2} x^k \text{d}F(x) \)?

Prop: Let \( C_k = \frac{2^k}{k!} \). Then \( m_k = C_k \) for \( k \geq 0 \).

Proof: Odd moments are 0 by symmetry. Integration by parts shows:
\[
m_{2k} = \frac{1}{2^{2k-1}} \int_{-2}^{2} x^{2k-1} x \text{d}F(x) = \frac{2k(2k-1)}{k!(2k+1)} m_{2k-2}.
\]
and even moments follow.

Remark: \( C_k \) is the \( k \)th Catalan number. They satisfy a recursion:
\[
C_k = \frac{k+1}{k+2} C_k \quad \text{where} \quad C_0 = 1.
\]

Proof: Write
\[
\int x^k \text{d}F(x) = \frac{1}{n} \sum_{i=1}^{n} x_i^k = \frac{1}{n} \sum_{i=1}^{n} \lambda_i (x_i)^k = \frac{1}{n} \text{Tr} X^k.
\]

Note: The \((i,j)\) entry of \( X^2 \) is
\[
\sum_{k=1}^{n} x_i x_j x_k x_l.
\]

The \((i,i)\) diagonal entry of \( X^k \) is
\[
\sum_{k=1}^{n} x_i k x_i x_{i+1} x_{i+2} x_i.
\]

Then
\[
\frac{1}{n} \text{Tr} X^k = \frac{1}{n} \sum_{i=1}^{n} x_i x_{i+1} x_{i+2} x_i.
\]

Each term in the sum is indexed by \((i, i+1, i+2)\). Associate to this a walk on the graph with vertices \( E_1, \ldots, E_n \) starting at \( E_i \) and traversing the edges \((i, i+1), (i+1, i+2), \ldots, (i, i+2)\).

\[
(3, 5, 2, 3, 8) \Rightarrow 3 \text{ select 2}
\]

The complete undirected graph on \( E_1, \ldots, E_n \) has \( \frac{n(n-1)}{2} \) distinct edges, including self-loops. These correspond to the \( \frac{n(n-1)}{2} \) independent entries of \( X \).
For any walk $(a_1, a_2, \ldots, a_n)$, if $E$ is the set of distinct edges traversed by this walk, and $k(e)$ is the number of times edge $e \in E$ is traversed, then by independence

$$E[X_{a_1}X_{a_2} \cdots X_{a_n}] = \prod_{e \in E} E[X_e^{k(e)}]$$

where $X_e = X_{a_j}$ if $e = a_1a_j$.

We divide the set of all walks $(a_1, a_2, \ldots, a_n)$ into three cases:

1. Some edge $e$ is traversed by this walk only once, i.e., $k(e) = 1$.

Then $E[X_{a_1}X_{a_2} \cdots X_{a_n}] = 0$, since $E[X_e^1] = 0$.

For all other walks, the number of distinct edges traversed is at most $k_e$.

- $\frac{k}{2}$ if $k_e$ is even.
- $\frac{k+1}{2}$ if $k_e$ is odd.

Then the number of distinct vertices visited is at most $k_e$, with $k_e$:

   - $\frac{k}{2}$ if $k$ is even, or $k-1=\frac{k-1}{2}$ if $k$ is odd.

2. The number of vertices visited is at most $\frac{k+1}{2}$.

How many such walks can there be? We can construct any such walk by first picking $\frac{k}{2}$ vertices from $E := \{e \mid k_e \geq 2\}$, then picking one at these vertices for each $e \in E$. So

$$\text{# such walks} \leq \binom{n}{\frac{k}{2}} \left(\frac{k}{2}\right)^k \leq \frac{k^k}{2^k} C_k$$

for a constant $C_k > 0$.

Recall $X_{i_{ij}} = \frac{1}{n} Z_{ij}$ where $Z_{ij}$ has bounded moments of all orders.

Then, since $E[X_e^{1/2}] = 1$, we have $E[X_e^{1/2}] \leq C_k$ for some constant $C_k > 0$, and any walk $W$.

Then

$$\left| E\left[ -\frac{1}{n} \sum_{W \text{ contains walk } X_{a_1}X_{a_2} \cdots X_{a_n}} \right] \right| \leq \frac{1}{n} \cdot \frac{1}{n^{1/2}} C_k \cdot \frac{1}{n^{1/2}} C_k \leq \frac{C_k}{\sqrt{n}}.

3. Each edge $e$ is traversed at least twice. If $k$ is even, and the number of visited vertices is exactly $\frac{k}{2} + 1$. Then:

   - Each edge is traversed exactly twice.

The graph of traversed edges is a tree of $\frac{k}{2}$ edges and no self-loops.

For each such walk, $k(e) \geq 2$ for $e$, and $E[X_e^2] = 0$, so

$$E[X_e^2] = \frac{1}{n} C_k.$$

To count the number of such walks: Call two walks $W_1$ and $W_2$ equivalent if some permutation of the labels $\{1, \ldots, n\}$ maps one to the other. This defines the set of such walks into equivalence classes, where walks in the same class have the same shape.

- Within each equivalence class, the walk is specified by the labels of the visited vertices in the order they are visited.

Then there are exactly $n(n-1)(n-2) \cdots (n-\frac{k}{2})$ walks per class.

Since the traversed graph is a tree, to specify an equivalence
class, we must specify for each of the \( k \) steps of the walk:

whether (a) we visit a new vertex, or (b) we return to its parent. This is in 1-1 correspondence with Dyck words of length \( k \), where:

\((a) \iff \text{(c)(1)(1)}\), \((b) \iff \text{(1)(c)(1)}\).

The number of equivalence classes is exactly the Catalan number \( C_k \). So

\[ E\left( \frac{1}{n} \sum \text{walks } X_{i_1} \ldots X_{i_n} \right) = \frac{1}{n} C_{k-1} n(n-1)(n-k) \frac{1}{n^k}. \]

Putting the three cases together:

\[ E\left( \frac{1}{n^k} \right) X^k = E\left[ \frac{1}{n} \sum \text{walks } \right] + E\left[ \frac{1}{n} \sum \text{walks } \right] + E\left[ \frac{1}{n} \sum \text{walks } \right] \]

\[ = 0 \quad \rightarrow \quad \text{as } n \rightarrow \infty \quad \rightarrow \quad \text{C.f. for } \text{over } k. \]

Thus for every fixed \( k \),

\[ E\left[ \sum x^k \mu(x) \right] = \int \left( \frac{1}{n} T x^k \right) \rightarrow m_k = \int x^k \mu(x). \]

Problem 1. Let \( X \in \mathbb{R}^n \) be symmetric, with \( X_{ij} = \frac{1}{2} Z_{ij} \) where:

- \( Z_{ij} = Z_{ji} \) and \( Z_{ij}, Z_{ji} \) are independent
- \( E[Z_{ij}] = 0 \) and \( E[Z_{ij}^2] < C_k \) for all \( ij \) and some constants \( C_k \).

There is a bounded, continuous function

\[ s : [0,1] \times [0,1] \rightarrow [0, \infty) \]

satisfying \( s(x,y) = s(y,x) \), so that

\[ E[Z_{ij}^2] = s\left( \frac{x_i - x_j}{n} \right). \]

Let \( \mu_n \) be the ESD of \( X \).

(a) Suppose, for each fixed \( y \), that \( s(y,x) dx = 1. \)

Show for every \( k \geq 0 \) that

\[ E\left[ \sum x^k \mu_n(x) \right] \rightarrow \int x^k \mu(x) \]

where \( \mu \) is still the semicircle law.

(b) More generally, for a Dyck word \( w \), consider all outward parentheses and let \( w_{-\infty} \), \( w_{+\infty} \) be the Dyck words contained in these parentheses. E.g.,

\[ w = \left( \left( \left( \right) \right) \left( \right) \right) \]

and

\[ w = \left( \left( \right) \left( \right) \left( \right) \right). \]

Define a function \( E(w,x) \) recursively by:
\[ \Phi(w;x) = \begin{cases} 1 & \text{if } w = \phi \\ \frac{1}{4} \sum_{i=1}^{4} s(x,y) \Phi(w_i,y)dy & \text{otherwise.} \end{cases} \]

Show for every even \( k > 0 \), that

\[ \mathbb{E} \left[ \int x^k d\mu(x) \right] \rightarrow \sum_{\text{dyadic w of length } k} \int \Phi(w,x) dx, \]

and that the limit is 0 for odd \( k \).