Scribe: Dantong Li
Lecturer: Tim Kunisky

Lecture #10: GOE and the Semicircle Law

1 Gaussian Orthogonal Ensemble (GOE)

Previously, we saw that the first moment method in the form of the Kac-Rice formula leads us to random matrix problems. In this lecture, we will talk about the random matrix theory and move toward free probability. We will start with the most critical random matrix theory example, the Gaussian Orthogonal Ensemble (GOE), which we usually denote by W distributed as:

$$\begin{aligned} W &\in \mathbb{R}^{N\times N}_{\text{sym}}, \\ W_{ij} = W_{ji} &\overset{iid}{\sim} \mathcal{N}(0,1), && \text{(for all } i<j) \\ W_{ii} &\overset{iid}{\sim} \mathcal{N}(0,2). \end{aligned}$$

We will not yet normalize W to have a spectrum with support independent of N, but rather will discover the correct normalization in our calculations.

The scaling of the diagonal entries' variances may be justified as follows. Defining an asymmetric Gaussian random matrix

$$\begin{aligned} G &\in \mathbb{R}^{N\times N}, \\ G_{ij} &\overset{iid}{\sim} \mathcal{N}(0,1), && \text{(for all } i,j) \end{aligned}$$

We have

$$W \overset{d}{=} \frac{G+G^T}{\sqrt{2}}$$

because on the diagonal we have the same variable added together and entries not on the diagonal is a sum of independent Gaussians. In particular, the important property of *rotational invariance* then follows:

$$W \overset{d}{=} QWQ^T \text{ for any deterministic orthognal matrix } Q \in O(N).$$

We prove this below in Proposition 1.1 and Corollary 1.2.

Proposition 1.1. *For $Q \in O(N)$, and a vector $g \sim \mathcal{N}(0, I_N)$, we have $g \overset{d}{=} Qg$.*

Proof. All the entries in Qg must be a linear combination of Gaussian variables, and a Gaussian vector is specified by its mean and covariance. Apparently, the mean of Qg is 0 by linearity. And its covariance is:

$$\mathbb{E}\left[(Qg)(Qg)^T\right] = Q\mathbb{E}\left[gg^T\right]Q^T = I.$$

□

Corollary 1.2. $QG \stackrel{d}{=} G$.

Proof. Apply Proposition 1.1 to each column of G. □

This W matrix is exactly what we saw in the last lecture up to normalization. Recall from the moment method lecture that what we are interested in from these matrices is their empirical spectral distribution (ESD), which is a random probability measure on $\mathbb{R}$:

$$\mu = \mu_W = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i}(W).$$

We want to understand what distribution this converges to when N goes to ∞:

$$\mu \to \rho,$$

for ρ some *deterministic* probability distribution.

The main task is to find what the moments of the ESD converge to:

$$\underbrace{\mathbb{E}\int x^k \, d\mu(x)}_{=:m_k} \to \int x^k \, d\rho(x)$$

So we will need to calculate these m_k's, which we do by expanding them as traces:

$$\begin{aligned} m_k &= \mathbb{E}\frac{1}{N}\sum_{i=1}^{N} \lambda_i^k \\ &= \frac{1}{N}\mathbb{E}\,\mathsf{Tr}\big(W^k\big) \\ &= \frac{1}{N}\mathbb{E}\sum_{i_1,\ldots,i_k \in [N]} W_{i_1 i_2} W_{i_2 i_3} \ldots W_{i_k i_1}. \end{aligned}$$

If we move the $\mathbb{E}$ inside the summation, to show what part of the summation has a non-zero contribution, we can think of each term of the summation as *a closed walk* on $1, \ldots, N$, as shown in Figure 1.

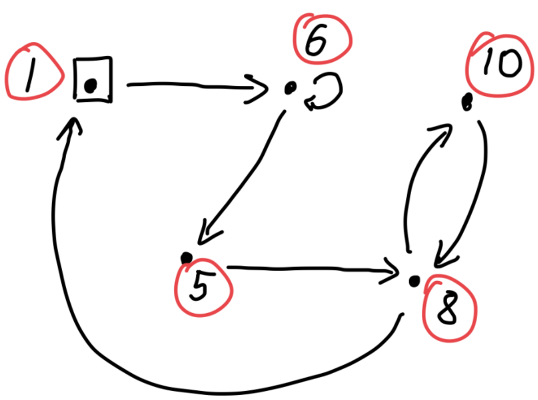

Figure 1: Closed Walk

Notice that to have a non-zero contribution, every edge suchmust be traversed an *even* number of times. If we ignore the labeling of the walk, we can define a *shape* to be a walk without vertex labels. *Good shapes* are those with non-zero contributions. Then

$$m_k = \frac{1}{N} \sum_{\text{good shapes}} \overbrace{f(\text{shape})}^{O_k(1)} \cdot \overbrace{N(N-1)\dots(N-\#\{\text{shape vertices}\}+1)}^{\text{number of such shapes} \approx N^{\#\{\text{shape vertices}\}}}.$$

Since we are looking at $N \to \infty$, the leading term of the summation is the shapes that maximize the *number of terms having such a shapes.* Let's consider the following two examples, as shown in Figure 2. k is the number of edges, for $k = 4$ and $k = 6$, we have two typical shapes. But the claim here is these are not the only good shapes that are the maximizers.

Remark 1.3. *if k is odd, then there are no good shapes, so $m_k = 0$.*

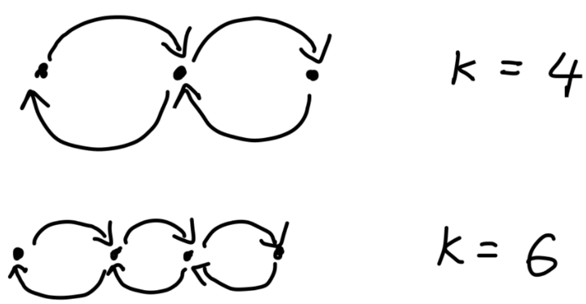

Figure 2: Shape Examples

Proposition 1.4. *$\{$Good shapes with k edges, $\frac{k}{2}+1$ vertices$\}$ = $\{$ "doubled" trees$\}$.*

Therefore,

$$m_k \approx N^{k/2} \times \# \left\{ \text{traversal-ordered trees on } \frac{k}{2}+1 \text{ vertices} \right\}$$

One can make a bijection that maps the elements of traversal-ordered trees in the above set to elements of *length-k balanced sequences of* { *"*`new`*", "*`back`*"*}, as illustrated in Figure 3.

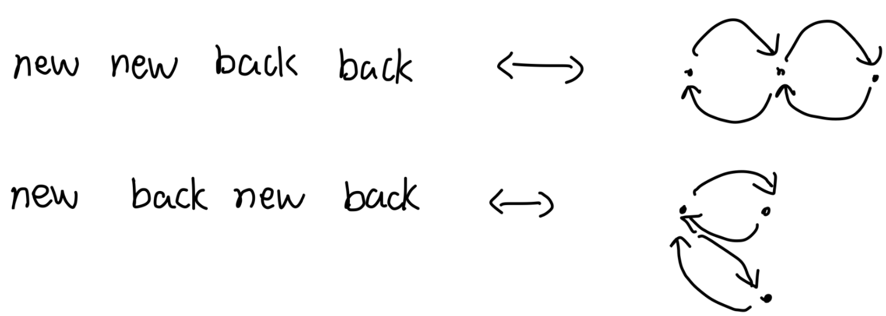

Figure 3: Bijection

Here we could understand "balanced" as, at any point, we can not have more `back`'s than `new`'s. This number also appears when *counting balanced parenthesizations* or *walks that never go below zero*, and is given by the *Catalan numbers*. These are

$$m_k = N^{k/2} C_{k/2} = N^{k/2} \frac{1}{k/2+1} \binom{k}{k/2}.$$

Theorem 1.5. *Let $\widehat{W} := \frac{1}{\sqrt{N}} W$. Then, for all $k \geq 1$,*

$$\mathbb{E} \frac{1}{N} \operatorname{Tr}\left(\widehat{W}^k\right) = \begin{cases} 0 & \text{if } k \text{ is odd,} \\ C_{k/2} & \text{if } k \text{ is even.} \end{cases}$$

2 The Semicircle Law

Now we know the moments of ρ, we want to invert them and find the probability distribution itself. And we will see the same device will also help calculate the formula for the Catalan number. This device is the *ordinary* moment generating function $M(x)$ defined as:

$$M(x) = \sum_{k \geq 0} m_k x^k.$$

Note that a valid sequence of length k should look like, for some $0 \leq i \leq k-2$,

$$\texttt{new} \underbrace{(\dots)}_{i} \texttt{back} \underbrace{(\dots)}_{k-2-i},$$

so we have the following recursion:

$$\begin{aligned} m_0 &= 1, \\ m_1 &= 0, \\ m_k &= \sum_{i=0}^{k-2} m_i m_{k-2-i} \text{ for } k \geq 2. \end{aligned}$$

Therefore,

$$\begin{aligned} M(x) &= \underbrace{1}_{k=0} + \underbrace{0 \cdot x}_{k=1} + \sum_{k \geq 2} \sum_{i=0}^{k-2} m_i m_{k-2-i} x^k \\ &= 1 + x^2 M(x)^2. \end{aligned}$$

Then we can solve the quadratic equation and get:

$$M(x) = \frac{1 - \sqrt{1-4x^2}}{2x^2}$$

Remark 2.1. *We choose $M(x) = \frac{1-\sqrt{1-4x^2}}{2x^2}$ rather than $M(x) = \frac{1+\sqrt{1-4x^2}}{2x^2}$ because we want $\lim_{x \to 0} M(x) \to 1 = m_0$.*

Remark 2.2. *We also see* $M(x) = \sum_{k\geq 0} C_k x^{2k}$*, since we only have even powers of* x *in* $M(x)$*. To get* C_k*, we compute the power series of* $(1+y)^{1/2} = \sum_{k\geq 0} \binom{1/2}{k} y^k$*, with the "formal binomial coefficients"* $\binom{1/2}{k} = \frac{\frac{1}{2}(\frac{1}{2}-1)\ldots(\frac{1}{2}-k+1)}{k!}$.

To recover the distribution, we can observe the following:

$$\begin{aligned} M(z) &= \sum_{k\geq 0} \mathop{\mathbb{E}}_{X\sim\rho}[X^k] z^k \\ &= \mathop{\mathbb{E}}_{X} \frac{1}{1-zX}. \end{aligned}$$

We instead consider the closely related Stieltjes transform, which is widely used in random matrix theory:

$$\begin{aligned} S(z) &= \mathop{\mathbb{E}}_{X} \frac{1}{X-z} \\ &= -\frac{1}{z} M(\frac{1}{z}) \\ &= -\frac{1}{z} \mathbb{E} \frac{1}{1-\frac{1}{z}X}. \end{aligned}$$

In our case, we may compute

$$S(z) = \frac{1}{2}\left(\sqrt{z^2-4} - z\right).$$

Remark 2.3. *All these transformations contain the same amount of information as the moment sequence.*

Now we will look into how to invert $S(z)$ to get the density function back. Note that if ρ has a density function $p(x)$, then

$$\begin{aligned} S(z) &= \mathop{\mathbb{E}}_{X} \frac{1}{X-z} \\ &= \int_{-\infty}^{\infty} \frac{p(x)}{X-z} dx. \end{aligned}$$

What we can do here is a complex analysis technique:

$$\begin{aligned} S(y+i\epsilon) - S(y-i\epsilon) &= \int_{-\infty}^{\infty} \frac{p(x)}{\underbrace{X-i\epsilon}_{\text{path}} - y} dx - \int_{-\infty}^{\infty} \frac{p(x)}{\underbrace{X+i\epsilon}_{\text{path}} - y} dx \\ &\approx \oint_C \frac{p(z)}{z-y} dx. \qquad \text{(roughly)} \end{aligned}$$

The contour C is illustrated in Figure 4.

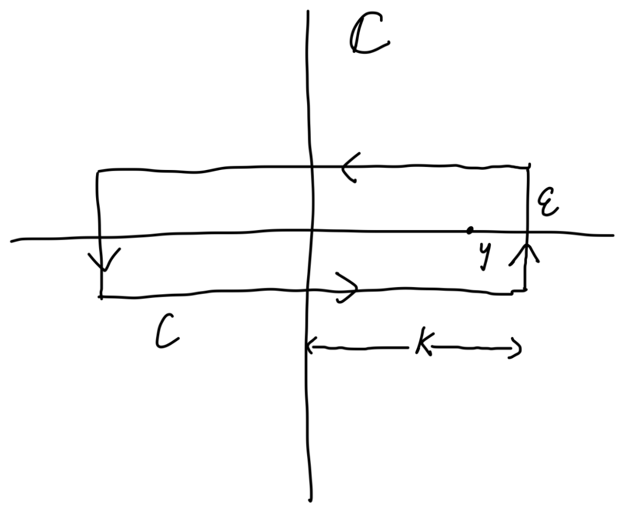

Figure 4: Integral Path

Proposition 2.4. *Under some additional conditions, by Cauchy's integral formula*

$$p(y) = \frac{1}{2\pi i} \lim_{\epsilon \to 0} \left\{ S(y + i\epsilon) - S(y - i\epsilon) \right\}.$$

For our $S(z)$, we need the "right" $\sqrt{z}$ that is compatible with our earlier calculations. The right choice is to avoid defining the square root of the negative real axis, as illustrated below in Figure 5.

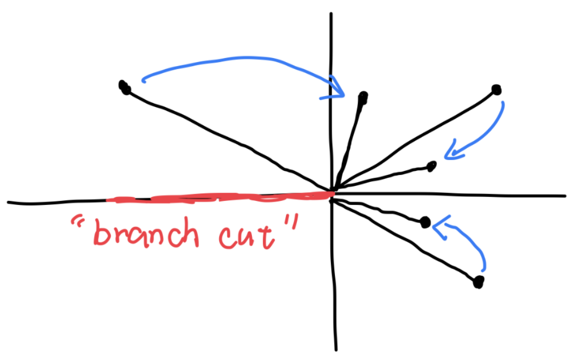

Figure 5: Square Root Function

Then we find, in our case,

$$\begin{aligned} p(y) &= \frac{1}{2\pi i} \lim_{\epsilon \to 0} \left\{ \frac{1}{2}\sqrt{(y+i\epsilon)^2 - 4} - \frac{1}{2}\sqrt{(y-i\epsilon)^2 - 4} \right\} \\ &= \frac{1}{2\pi i} \begin{cases} 0 & \text{if } y^2 \geq 4, \\ i \cdot \frac{1}{2}\sqrt{y^2 - 4} \cdot 2 & \text{if } y^2 < 4 \end{cases} \\ &= \frac{1}{2\pi} \begin{cases} 0 & y^2 \geq 4, \\ \sqrt{y^2 - 4} & y^2 < 4. \end{cases} \end{aligned}$$

And this is exactly the density of the semicircle law.

Theorem 2.5. *The ESD of the normalized GOE matrix* $\widehat{W}$ *converges in moments to the* semicircle law *having density*

$$p(y) = \frac{1}{2\pi}\sqrt{y^2 - 4} \cdot \mathbb{1}\left\{y \in [-2, 2]\right\}.$$

Remark 2.6. *All of these calculations depend very little on the distribution of the entries of* W*. Let* W_{ii} *be anything reasonable,* $\mathbb{E}W_{ij} = 0$ *and* $\mathbb{E}w_{ij}^2 = 1$*, and a small amount of further regularty, then we have similar decay in the higher moments of the entries so that we could carry through the "tree counting" arguments we did previously, then the same exact argument is applicable to any such random matrix to get the semicircle law in the end.*

Next time we will look back and see how we can use this argument to handle more complicated matrices resulting from the Kac-Rice formula.