

Lectures 2 – 3 : Wigner's semicircle law

Notes prepared by: M. Koyama

As we set up last week, let $M_n = [X_{ij}]_{i,j=1}^n$ be a symmetric $n \times n$ matrix with Random entries such that

- $X_{i,j} = X_{j,i}$
- $X_{i,j}$ s are *iid* for all $i < j$, and X_{jj} are iid for all j with

$$E[X_{ij}^2] = 1, \quad E[X_{ij}] = 0$$

- All moments exists for each entries.

We considered the eigenvector of this random matrix;

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$$

which turns out to be random elements depending continuously on M_n ;

Lemma 1. *If \mathcal{H}_n is a topological space of $n \times n$ matrix with topology derived from the usual metric on product Lebesgue measurable space, then $\lambda_i(\mathcal{H})$ is a continuous function on \mathcal{H}_n .*

Proof. Let $H = [h_{ij}]_{i,j=1}^n$ be an element in \mathcal{H}_n . We know that

$$\|H\|_k = \sqrt[k]{\text{Tr}(H^k)} = \sqrt[k]{\sum \lambda_i^k}$$

So for example, $\|H\|_2 = \sqrt{\sum_i \lambda_i^2}$. Note that therefore $\|H\|_2 \geq \max(\lambda_n, -\lambda_1)$. Our goal is to obtain λ in terms of H . So it is good if we can say

$$\lim_{k \rightarrow \infty} \|H\|_k \rightarrow \lambda_n$$

because λ_n dominates all the other eigen vectors, maybe except λ_1 . Clearly, this logic might not work because of the presence of negative eigen values including λ_1 . To fix this problem we may just shift the matrix by $\|H\|$. In particular, we can claim

$$\lim_{k \rightarrow \infty} \sqrt[k]{\text{Tr}((H + \|H\|I)^k)} \rightarrow \lambda_n + \|H\|$$

To be more precise,

$$\lambda_n(H) + \|H\| \leq \sqrt[k]{\text{Tr}((H + \|H\|I)^k)} \quad (1)$$

$$\leq \sqrt[k]{n}(\lambda_n(H) + \|H\|) \quad (2)$$

$$\leq \lambda_n(H) + \|H\|. \quad (3)$$

Having obtained $\lambda_n \dots \lambda_k$, we can inductively obtain the λ_{k-1} by simply taking the limit of

$$\sqrt[k]{\text{Tr}((H + \|H\|I)^k) - \sum_{i=1}^k (\lambda_n(H) + \|H\|)^k}$$

□

This allows us to induce the random measure

$$\nu_n = \frac{1}{n} \sum \delta_{\frac{\lambda_i}{\sqrt{n}}}.$$

The Wigner's semicircle law claims that this ν_n has a nice distributional limit.

Theorem 2.

$$\frac{1}{n} \sum \delta_{\frac{\lambda_i}{\sqrt{n}}} \Rightarrow \nu$$

where $\frac{\nu}{dx} = \frac{1}{2\pi} \sqrt{4 - x^2} 1_{(|x| \leq 2)}$.

We will use Borrel Cantelli lemma and Carleman's condition for moment problem to show this fact. Consider the random variable

$$X_{n,k} = \int x^k d\nu_n.$$

We will show

1.

$$\mathbf{E}X_{n,k} \rightarrow \mathbf{c}_k = \int \mathbf{x}^k d\nu \quad (4)$$

2.

$$\mathbf{Var}(X_{n,k}) \leq \frac{\mathbf{c}_k}{n^2} \quad (5)$$

How do they help? Suppose these two statements are true. Then we can use Borrel Cantelli lemma to show that

$$P(|X_{n,k} - EX_{n,k}| > \frac{1}{\sqrt[4]{n}}) \leq E(X_{n,k} - EX_{n,k})^2 \sqrt{n} \quad (6)$$

$$= \text{Var}(X_{n,k}) \sqrt{n} \quad (7)$$

$$= O(1/n^{3/2}) \quad (8)$$

Thus $P(|X_{n,k} - EX_{n,k}| > \frac{1}{\sqrt[3]{n}} \text{ i.o.}) = 0$ and $|X_{n,k} - EX_{n,k}| < \frac{1}{\sqrt[3]{n}}$ for some large n almost surely. If this is the case, then ν_n can be shown to be tight because this means $X_{n,k}$ is bounded by some constant C and hence by Chebyshev

$$\nu_n(\{x : |x| > m\}) < \frac{C}{m^k}.$$

We can therefore choose a converging subsequence $\nu_{n(j)}$ of measures that converge to ν^* . We would now like to show that any of these subsequential limits ν^* of converging subsequences equals to ν . In this way, we can establish that any subsequence $\nu_{n(k)}$ has further subsequence that converges to ν . This can be done if we can characterize ν by its moments, because we know that ν^* 's moments for all subsequence agrees by the claim (1). This can be done using the following useful criterion.

Theorem 3. (*Carleman's condition:*) Suppose

$$\sum_{k=1}^{\infty} \frac{1}{\mu_k^{1/2k}} = \infty$$

. Then there is at most one measure F such that $\int x^k dF(x) = \mu_k$ for all positive integer k . This criterion can be made stronger: in fact, the conclusion above holds if

$$\limsup \frac{\mu_{2k}^{1/2k}}{2k} = r < \infty.$$

The logic behind the proof of this claim follows from the fact that the characteristic function $E[\exp(iXt)]$ characterizes the distribution of X . We can consider the Taylor polynomial of $\exp(iXt)$. If $E[\exp(iXt)] = \sum \frac{(it)^k E[X^k]}{k!}$, then the moment indeed determines the characteristic function.

Let's hence check if the Carleman's condition applies to our case. Put

$$c_k = \int_{-2}^2 \frac{1}{2\pi} x^k \sqrt{4-x^2} dx.$$

If k is odd, then $c_k = 0$. Therefore put $k = 2n$. Then

$$c_k = \frac{1}{\pi} \int_{[0,2]} x^{2n} \sqrt{4-x^2} dx \tag{9}$$

$$= \frac{1}{\pi} \int_{[0,\pi/2]} \sin^{2n}(t) \cos^2(t) 2^{2n+2} dt \tag{10}$$

$$= \frac{1}{\pi} \int_{[0,\pi/2]} 2^{2n+2} (\sin^{2n}(t) - \sin^{2n+2}(t)) dt \tag{11}$$

$$= \frac{1}{\pi} 2^{2n+2} \frac{(2n)!}{n! 2^{2n}} \frac{\pi}{2} \left(1 - \frac{(2n+2)(2n+1)}{4(n+1)^2} \right) \tag{12}$$

$$= \binom{2n}{n} \frac{1}{n+1} < 4^n \tag{13}$$

We used the fact

$$\int_{[0,\pi/2]} \sin^{2\ell}(t) dt = \frac{(2\ell)!}{(\ell!)^2 2^{2\ell}} \frac{\pi}{2}$$

Therefore $\frac{\mu_{2k}^{1/2k}}{2k} < \frac{(4^{k/2})^{1/2k}}{2k} = \frac{\sqrt[4]{4}}{2k}$ and the claim follows.

Therefore, it remains to show (1) and (2) in (0.4) and (0.5).

Proof of (1)

Let us begin with (1). We will achieve this by a way of "controlled brute force". Note that

$$E \int x^k d\nu_n = E \frac{1}{n} \sum \left(\frac{\lambda_i}{\sqrt{n}} \right)^k \quad (14)$$

$$= n^{-1-\frac{k}{2}} E(\text{Tr} M_n^k) \quad (15)$$

$$= n^{-1-\frac{k}{2}} \sum E(X_{i_1, i_2} X_{i_2, i_3} X_{i_3, i_4} \cdots X_{i_k, i_1}) \quad (16)$$

To organize this, whenever we have k -tuple $(i_1, i_2, \dots, i_k) = I$, put

$$E(I) = E(X_{i_1, i_2} X_{i_2, i_3} X_{i_3, i_4} \cdots X_{i_k, i_1}).$$

First, observe that $E(I)$ is bounded by some constant B_k . This can be seen by applying Cauchy Shwartz inequality inductively.

Let us represent each I by a directed closed path with vertices $\{1, 2, 3, \dots, n\} = V(I)$ and edges $\xi(I) = \{(i_a, i_{a+1}); a = 1, \dots, k, i_{k+1} = i_1\}$. For example, if $I = (2, 3, 1, 2, 2, 1)$ then this will correspond to the directed adjacency matrix ¹

$$\begin{bmatrix} 0 & 2 & 0 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad (17)$$

Now, **skeleton** of a directed graph is a undirected graph induced by the directed graph by replacing all the multiedges by edges. For example, the skeleton of the graph above is given by the adjacency matrix ²

$$\begin{bmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}. \quad (18)$$

¹entries in the a_{ij} represents the number of edges from vertex i to vertex j .

²entries in the a_{ij} represents the number of edges between vertex i to vertex j .

Here, remark that that $E[I] = 0$ unless every edge in the skeleton is used at least twice. If, for example, an edge (i, j) happens only once, then

$$E[I] = E[X_{i,j}]E \left[\prod_{e \in \xi(I) \setminus \{(i,j)\}} X_e \right] = 0$$

This implies that if

$$\mathbf{E}(\mathbf{I}) \neq \mathbf{0} \quad \text{then} \quad \xi(\mathbf{I}) \leq \frac{\mathbf{k}}{2}.$$

This bound let us also put a bound on $V(I)$;

Lemma 4. *Given any graph G , denote the vertex set by $V(G)$ and edge set by $E(G)$. Then $|V(G)| \leq |E(G)| + 1$.*

Proof. To see this, first assume that G is a tree. Note that removing a leaf from the graph removes one edge and one vertex. We may continue removing leaves from the Graph until K_2 (complete graph of 2 vertices) remains. Removing a leaf from K_2 results in K_1 . Thus $V(G) = E(G) + 1$ in this case. For a generic graph G , we may remove edges from the graph until we obtain its spanning tree G'' . If we removed m edges in this process, then $V(G) = E(G') + 1 + m$ and the claim follows. \square

Thus, we have

$$\mathbf{E}(\mathbf{I}) \neq \mathbf{0} \quad \text{then} \quad \mathbf{V}(\mathbf{I}) \leq \lfloor \frac{\mathbf{k}}{2} \rfloor + 1.$$

We are now in position to bound the expectation of $X_{n,k}$.

Lemma 5.

$$\left| E \left[\int x^k d\nu_n \right] \right| \leq \frac{c_k}{\sqrt{n}}$$

Proof.

$$E \int x^k d\nu_n = \frac{1}{n^{k/2+1}} \sum_I E(I) \tag{19}$$

$$= \sum_{V(I) \leq \lfloor \frac{k}{2} \rfloor + 1} E(I) \tag{20}$$

$$\leq \frac{B_k}{n^{1+k/2}} \left| \{I; V(I) \leq \lfloor \frac{k}{2} \rfloor + 1\} \right| \tag{21}$$

Temporarily, consider $V(I) = \ell$ for a fixed ℓ . How many ways can we choose I ? Most naive bound on this number is indeed $n^\ell * \ell^k$. It turns out that this naive bound suffices. From the inequality that we obtained above, we see that if $\ell < \frac{k}{2} + 1$ then the terms with $V(I) = \ell$ will vanish in limit. Thus we can ignore the odd k all together in the limit. Let us therefore consider the case of even

k . When k is even, we see that $V(I) \leq \frac{k}{2} + 1$. If the inequality is strict, again $E \int x^k d\nu_n \rightarrow 0$ in the limit. **Therefore, asymptotically, we can restrict our case to when $V(I) = \frac{k}{2} + 1$ and $\xi(I) \leq \frac{k}{2}$.** Because $V(I) \leq \xi(I) + 1$, we have $\xi(I) = \frac{k}{2}$ necessarily. We are thus considering directed graphs for which the skeletons are trees, and there are exactly two edges between two adjacent vertices. This kind of directed graph is called a **double tree**. Below is a directed adjacency matrix for an example of a double tree;

$$\begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad (22)$$

Now, if I is a double tree, clearly

$$E(I) = E \left(\prod_{e \in \xi(I)} X_e^2 \right) = \prod_{e \in \xi(I)} E(X_e^2) = 1$$

Thus, all together, we obtain the following statement;

Proposition 6.

$$\lim_{n \rightarrow \infty} E \left(\int x^k d\nu_n \right) = \lim_{n \rightarrow \infty} \frac{1}{n^{1+\frac{k}{2}}} * (\text{Number of double trees with } n \text{ vertices})$$

Our proof of (1) is therefore simplified to the counting of the number of double trees with n vertices. To answer this, first let us answer the following question; "If I fix a shape of a tree, just how many double trees of that shape exist?" We may achieve this by making a bijection between the shape of a double tree and a random walk on \mathbb{N} beginning from 0 and returning in exactly k -step. For example, suppose that a double tree is given by the directed adjacency matrix above; then fixing the vertex 1 as the starting point of the walk, the shape of this double tree corresponds to the random walk $(0, 1, 2, 1, 2, 1, 0)$ (the k th entry is the distance of the walker from the vertex 1 at k th step). Counting this way, we will show next week that the shape of a double tree with $\frac{k}{2}$ edges are given by

$$\binom{k}{\frac{k}{2}} \frac{1}{k+1}.$$

Now, given a fixed shape, the number of double trees of that shape is given by

$$\underbrace{\binom{n}{\frac{k}{2} + 1}}_{\text{choosing the vertices}} \underbrace{\left(\frac{k}{2} + 1 \right)!}_{\text{permutation}}$$

Thus at last, we obtain that

$$\lim_{n \rightarrow \infty} \int x^k d\nu_n = \lim_{n \rightarrow \infty} \frac{1}{n^{k/2+1}} \binom{k}{\frac{k}{2}} \frac{1}{k+1} \frac{1}{\frac{k}{2}+1} n(n-1) \cdots \left(n - \frac{k}{2} \right) \quad (23)$$

$$= \binom{k}{\frac{k}{2}} \frac{1}{\frac{k}{2}+1} = \binom{2n}{n} \frac{1}{n+1} \quad (24)$$

and the claim follows. \square