Wigner's Semicircular Theorems

October 16, 2024

1 Definitions

Assume we have two families of independent and identical real-valued random variables with zero-mean $\{Z_{i,j}\}_{1\leq i<j}$ and $\{Y_i\}_{1\leq i}$ such that $\mathbb{E}[Z_{1,2}^2] = 1 = \mathbb{V}\mathrm{ar}[Z_{1,2}]$ for all integers $k \geq 1$.We also have,

$$r_k := \max\left\{\mathbb{E}[|Z_{1,2}^k|], \mathbb{E}[|Y_1^k|]\right\} < \infty$$

Then we define Wigner matrices as the symmetric $N \times N$ matrices X_N as

$$X_N(i,j) = X_N(j,i) = \begin{cases} \frac{Z_{i,j}}{\sqrt{N}} & \text{if } i < j, \\ \frac{Y_i}{\sqrt{N}} & \text{if } i = j. \end{cases}$$

Let λ_i^N be the real ordered eigenvalues of X_N and we define the empirical distribution of the eigenvalues as the probability measure on $\mathbb{R}$ defined as

$$L_N = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i^N}$$

Denote the semicircle distribution as the probability distribution $\sigma(x)dx$ on $\mathbb{R}$ with density

$$\sigma(x) = \frac{1}{2\pi}\sqrt{4-x^2}\mathbf{1}_{\{|x|\leq 2\}}$$

2 Statement of Theorem

Wigner's Semicircle Theorem for Real Matrices:
For all Wigner matrices, the empirical measure L_N converges in probability to the semicircle distribution
i.e. for all $f \in C_b(\mathbb{R})$ and for all $\epsilon > 0$

$$\lim_{N\to\infty} \mathbb{P}\left(|\langle L_N, f\rangle - \langle \sigma, f\rangle| > \epsilon\right)$$

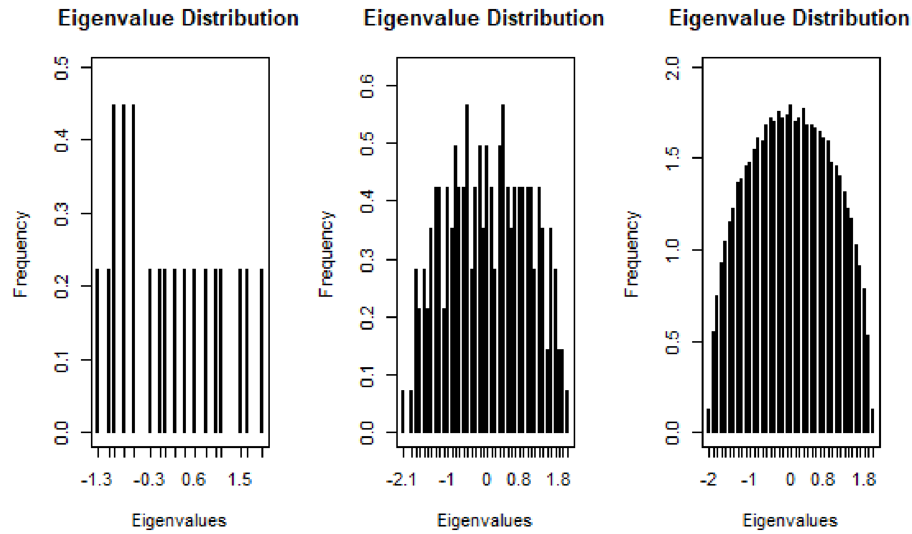

Figure 1: Three realisations of the eigenvalue distribution for $N = 20, 200, 3000$

3 Helpful Lemmas

Lemma 1:
For all $k \in \mathbb{N}$,

$$\lim_{N\to\infty} \langle \overline{L}_N, x^k \rangle = \langle \sigma, x^k \rangle$$

We denote $m_k := \langle \sigma, x^k \rangle$ and $m_k^N = \langle \overline{L}_N, x^k \rangle$

Proof of Lemma 1:
We start with basic matrix algebra

$$\begin{aligned}
\langle \overline{L}_N, x^k \rangle &= \frac{1}{N}\mathbb{E}[\mathrm{tr}(X_N^k)] \\
&= \frac{1}{N} \sum_{i_1,\ldots,i_k=1}^{N} \mathbb{E}[X_N(i_1,i_2)X_N(i_2,i_3)\cdots X_N(i_{k-1},i_k)X_N(i_k,i_1)] \\
&:= \frac{1}{N} \sum_{i_1,\ldots,i_k=1}^{N} \mathbb{E}[T_{\mathbf{i}}^N] := \frac{1}{N} \sum_{i_1,\ldots,i_k=1}^{N} \overline{T}_{\mathbf{i}}^N \qquad \text{where } \mathbf{i} = (i_1, ..., i_k) \qquad (3.1)
\end{aligned}$$

We now transfer to the language of combinatorics to complete this proof...

Given a set $\mathcal{J}$, a $\mathcal{J}$-letter is an element of $\mathcal{J}$, and a $\mathcal{J}$-word is a finite sequence of $\mathcal{J}$-letters, at least one letter long. A $\mathcal{J}$-word is closed if the first and last letters are the same, and two $\mathcal{J}$-words are equivalent, w_1 w_2, if there exists a bijection on $\mathcal{J}$ mapping one to the other. If $\mathcal{J} = \{1, ..., N\}$, we call it an N-word. The length of a word w, $l(w)$, is the number of letters in w, its weight, $\mathrm{wt}(w)$, is the number of distinct letters in the word, and the support of w, $\mathrm{supp}(w)$, is the set of letters in w.

We can then associate to each word $w = s_1, ..., s_k$ a graph $G_w = (V_w, E_w)$ where vertices $V_w = \mathrm{supp}(w)$, and edges $E_w = \{\{s_i, s_{i+1}\}; i = 1, ..., k-1\}$. We call the set of self edges $E_w^s = \{e \in E_w : e = \{u, u\}, u \in V_w\}$, and the set of connecting edges $E_w^c = E_w \backslash E_w^s$. Note that G_w is connected as w defines a path connecting all the vertices, and defines a cycle if the word is closed. Denote N_e^w for $e \in E_w$ the number of times the path w traverses edge e.

Now note that any k-tuple of integers $\mathbf{i}$ defines a closed word $w_{\mathbf{i}} = i_1 i_2 .... i_k i_1$ with length $k+1$. Then,

$$\overline{T}_{\mathbf{i}}^N = \frac{1}{N^{\frac{k}{2}}} \prod_{e \in E_{w_{\mathbf{i}}}^c} \mathbb{E}[Z_{1,2}^{N_e^{w_{\mathbf{i}}}}] \prod_{e \in E_{w_{\mathbf{i}}}^s} \mathbb{E}[Y_1^{N_e^{w_{\mathbf{i}}}}] \qquad (3.2)$$

Hence we see that $\overline{T}_{\mathbf{i}}^N = 0$ unless $N_e^{w_{\mathbf{i}}} \geq 2$ for all $e \in E_{w_{\mathbf{i}}}$. Furthermore, $\mathrm{wt}(w_{\mathbf{i}}) \leq \frac{k}{2} + 1$

Note that if $N \geq t$, then there are $C_{N,t} = N(N-1)...(N-t+1)$ N-words equivalent to a given N-word with weight t.

We define $\mathscr{W}_{k,t}$ to be the set of representatives for equivalency classes of closed N-words w with length $k+1$, weight t, and with $N_e^w \geq 2$ for all $e \in E_w$. We then deduce from (3.1) and (3.2) that

$$\langle \overline{L}_N, x^k \rangle = \sum_{t=1}^{\lfloor \frac{k}{2} \rfloor + 1} \frac{C_{N,t}}{N^{\frac{k}{2}+1}} \sum_{w \in \mathscr{W}_{k,t}} \prod_{e \in E_w^c} \mathbb{E}[Z_{1,2}^{N_e^w}] \prod_{e \in E_w^s} \mathbb{E}[Y_1^{N_e^w}] \qquad (3.3)$$

The cardinality of $\mathscr{W}_{k,t}$ is bounded by the number of closed $\mathcal{J}$-words of length $k+1$ when the cardinality of $\mathcal{J}$ is $t \leq k$

$$|\mathscr{W}_{k,t}| \leq t^k \leq k^k$$

Thus (3.3) and the finiteness of r_k imply

$$\lim_{N \to \infty} \langle \overline{L}_N, x^k \rangle = \begin{cases} 0 & \text{if } k\text{odd,} \\ \sum_{w \in \mathscr{W}_{k, \frac{k}{2}+1}} \prod_{e \in E_w^c} \mathbb{E}[Z_{1,2}^{N_e^w}] \prod_{e \in E_w^s} \mathbb{E}[Y_1^{N_e^w}] & \text{if } k\text{even.} \end{cases}$$

since

$$\lim_{N \to \infty} \sum_{t=1}^{\lfloor \frac{k}{2} \rfloor + 1} \frac{C_{N,t}}{N^{\frac{k}{2}+1}} \approx \sum_{t=1}^{\lfloor \frac{k}{2} \rfloor + 1} N^{t - \frac{k}{2} - 1}$$

We define a closed word w of length $k+1 \geq 1$ is a Wigner word if $k=0$, or k even and w equivalent to an element of $\mathscr{W}_{k,\frac{k}{2}+1}$

Note that if $w \in \mathscr{W}_{k,\frac{k}{2}+1}$, then G_w is a tree, and also $E_w^s = \emptyset$. This implies that

$$\lim_{N\to\infty} \langle \overline{L}_N, x^k \rangle = |\mathscr{W}_{k,\frac{k}{2}+1}| \quad \text{for k even}$$

Now let k be even and choose $\mathscr{W}_{k,\frac{k}{2}+1}$ such that all words $w = v_1, ..., v_{k+1}$ satisfies for $i = 1, ..., k+1$, that $\{v_1, ..., v_i\}$ is an interval in the integers, beginning with 1. Note each element $w \in \mathscr{W}_{k,\frac{k}{2}+1}$ determines a path $v_1, v_2, ..., v_k, v_{k+1} = v_1$ length $k+1$ on a tree G_w called the exploration process for w.

Let $d(v, v')$ denote the distance between v, v' on a tree G_w which is the length of the shortest path between them. Then we set $x_i = d(v_{i+1}, v_1)$ and see that each word $w \in \mathscr{W}_{k,\frac{k}{2}+1}$ defines a Dyck path $D(w) = (x_1, x_2, ..., x_k)$ of length k.

Conversely, given a Dyck path $x = (x_1, ..., x_k)$ can construct a word $w = T(x) \in \mathscr{W}_{k,\frac{k}{2}+1}$ by recursively constructing an increasing sequence $w_2, ..., w_k = w$ of words as follows:

1) $w_2 = (1, 2)$

2) For $i > 2$, if $x_{i-1} = x_{i-2} + 1$, then w_i is obtained by adjoining on the right of w_{i-1} the smallest positive integer not in w_{i-1}. Otherwise w_i is obtained by adjoining on the right of w_{i-1} next to the last letter of w_{i-1}

Note that for all i, G_{w_i} is a tree since G_{w_2} is a tree and at stage i we either add a leaf, or a backtrack is added. Hence the distance in G_{w_i} between the first and last letter of w_i is x_{i-1}, and so $D(w) = (x_1, ..., x_k)$.

With our choice of representation, $T(D(w)) = w$ as each 'up' in $D(w)$ starting at location $i-2$ corresponds to the adjoinment on the right of w_{i-1} of a new letter, uniquely determined by $\text{supp}(w_{i-1})$, but each 'down' at $i-2$ corresponds to the joining of penultimate letter in w_{i-1}. Hence there is a bijection between Dyck paths of length k and $\mathscr{W}_{k,\frac{k}{2}+1}$.

Now we will calculate how many unique Dyck paths of length $2k$ exist for each k. First let B_k be the number of Bernoulli walks S_n length $2k$ such that $S_{2k} = 0$, and let $\overline{B}_k$ be the number of Bernoulli walks S_n length $2k$ such that $S_{2k} = 0$ and $S_t < 0$ for some $t < 2k$. Hence the number of Dyck paths is $B_k - \overline{B}_k$.

Now $B_k = \binom{2k}{k}$ as we must have k 'up' steps and k 'down' steps. Then $\overline{B}_k = \binom{2k}{k-1}$ as must have $k+1$ 'down' steps and $k-1$ 'up' steps after the first time going negative. Hence the number of Dyck paths is $\binom{2k}{k} - \binom{2k}{k-1} = C_k$

where C_k is the k^{th} Catalan number.

We finish by showing that $m_{2k} = C_k$ for all integers k. First note that $m_{2k+1} = 0$ by symmetry over the integral. Now,

$$\begin{aligned}
m_{2k} &= \int_{\mathbb{R}} x^{2k}\sigma(x)\,dx \\
&= \frac{1}{2\pi}\int_{-2}^{2} x^{2k}\sqrt{4-x^2}\,dx \\
&= \frac{1}{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (2sin(\theta))^{2k}\sqrt{4-4sin^2(\theta)}2cos(\theta)\,d\theta \\
&= \frac{2^{k+1}}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} sin^{2k}(\theta)cos^2(\theta)\,d\theta \\
&= \frac{2^{k+1}}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} sin^{2k}(\theta)(1-sin^2(\theta))\,d\theta \\
&= \frac{2^{k+1}}{\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} sin^{2k}(\theta)\,d\theta - (2k+1)m_{2k} \qquad \text{from integrating by parts}
\end{aligned}$$

Hence $m_{2k} = \frac{2^{k+1}}{\pi(2k+2)}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} sin^{2k}(\theta)\,d\theta = \frac{\binom{2k}{k}}{k+1} = C_k$ if fully evaluated.

So to conclude, we have shown

$$\lim_{N\to\infty}\langle \overline{L}_N, x^k\rangle = |\mathscr{W}_{k,\frac{k}{2}+1}| = C_{\frac{k}{2}} = m_k = \langle \sigma, x^k\rangle$$

and we are done. $\square$

Lemma 2:
For all $k \in \mathbb{N}$ and $\epsilon > 0$,

$$\lim_{N\to\infty}\mathbb{P}\left(\left|\langle L_N, x^k\rangle - \langle \overline{L}_N, x^k\rangle\right| > \epsilon\right) = 0$$

Proof of Lemma 2:
By Chebyshev's Inequality, we can equivalently prove

$$\lim_{N\to\infty}|\mathbb{E}[\langle L_N, x^k\rangle^2] - \langle \overline{L}_N, x^k\rangle^2| = 0$$

As similar to previous (3.1),

$$\mathbb{E}[\langle L_N, x^k\rangle^2] - \langle \overline{L}_N, x^k\rangle^2 = \frac{1}{N^2}\sum_{\substack{i_1,\ldots,i_k=1\\ i'_1,\ldots,i'_k=1}}^{N} \overline{T}^N_{\mathbf{i},\mathbf{i}'} \qquad (3.4)$$

where $\overline{T}^N_{\mathbf{i},\mathbf{i'}} = \mathbb{E}[T^N_{\mathbf{i}} T^N_{\mathbf{i'}}] - \mathbb{E}[T^N_{\mathbf{i}}]\mathbb{E}[T^N_{\mathbf{i'}}]$

Given a set $\mathcal{J}$, a $\mathcal{J}$-sentence is a finite sequence of $\mathcal{J}$-words at least one word long. Two sentences are equivalent, a_1 a_2, if there is a bijection on $\mathcal{J}$ mapping one to the other. For a sentence $a = (w_1, w_2, ..., w_n)$, the support supp(a) $= \bigcup_{i=1}^n \text{supp}(w_i)$, and the weight, $wt(a)$, is the cardinality of supp(a).

We can assign graphs to sentences as previous. Given a sentence $a = (w_1, ..., w_k) w_i = s^i_1, s^i_2, ..., s^i_{l(w)}$ we set $G_a = (V_a, E_a)$ with vertices $V_a = \text{supp}(a)$ and edges $E_a = \left\{\left\{s^i_j, s^i_{j+1}\right\}; j = 1, ..., l(w_i) - 1, i = 1, ..., k\right\}$. We define the set of self edges and connecting edges in the same way as before, but note the graphs may be disconnected this time and sentence a defines k paths G_a. Define N^a_e the number of times the union of the paths traverse edge e. Equivalent sentences generate the same graph and passage counts.

We recall closed words $w_{\mathbf{i}}, w_{\mathbf{i'}}$ of length $k+1$ and define the two word sentence $a_{\mathbf{i},\mathbf{i'}} = (w_{\mathbf{i}}, \mathbf{i'})$, then

$$\begin{aligned}\overline{T}^N_{\mathbf{i},\mathbf{i'}} = \frac{1}{N^k}\Big(&\prod_{e \in E^c_{a_{\mathbf{i},\mathbf{i'}}}} \mathbb{E}[Z_{1,2}^{N_e^{a_{\mathbf{i},\mathbf{i'}}}}] \prod_{e \in E^s_{a_{\mathbf{i},\mathbf{i'}}}} \mathbb{E}[Y_1^{N_e^{a_{\mathbf{i},\mathbf{i'}}}}] \\ &- \prod_{e \in E^c_{w_{\mathbf{i}}}} \mathbb{E}[Z_{1,2}^{N_e^{w_{\mathbf{i}}}}] \prod_{e \in E^s_{w_{\mathbf{i}}}} \mathbb{E}[Y_1^{N_e^{w_{\mathbf{i}}}}] \prod_{e \in E^c_{w_{\mathbf{i'}}}} \mathbb{E}[Z_{1,2}^{N_e^{w_{\mathbf{i'}}}}] \prod_{e \in E^s_{w_{\mathbf{i'}}}} \mathbb{E}[Y_1^{N_e^{w_{\mathbf{i'}}}}]\Big)\end{aligned} \tag{3.5}$$

Hence $\overline{T}^N_{\mathbf{i},\mathbf{i'}} = 0$ unless $N_e^{a_{\mathbf{i},\mathbf{i'}}} \geq 2$ for all edges, and $E_{w_{\mathbf{i}}} \cap E_{w_{\mathbf{i'}}} \neq \emptyset$

Again if $N \geq t$, there are $C_{N,t}$ N-sentences equivalent to a given N-sentence weight t

Let $\mathscr{W}^{(2)}_{k,t}$ denote the set of representatives for equivalence classes of sentences a with weight t comprised of 2 closed t-words (w_1, w_2) each with length $k+1$ with $N^a_e \geq 2$ for all $e \in E_a$ and $E_{w_{\mathbf{i}}} \cap E_{w_{\mathbf{i'}}} \neq \emptyset$. Then from (3.4) and (3.5),

$$\begin{aligned}\mathbb{E}[\langle L_N, x^k\rangle^2] - \langle \overline{L}_N, x^k\rangle^2 = \sum_{t=1}^{2k} \frac{C_{N,t}}{N^{k+2}} \sum_{a=(w_1,w_2)\in\mathscr{W}^{(2)}_{k,t}} \Big(&\prod_{e \in E^c_{a_{\mathbf{i},\mathbf{i'}}}} \mathbb{E}[Z_{1,2}^{N_e^{a_{\mathbf{i},\mathbf{i'}}}}] \prod_{e \in E^s_{a_{\mathbf{i},\mathbf{i'}}}} \mathbb{E}[Y_1^{N_e^{a_{\mathbf{i},\mathbf{i'}}}}] \\ &- \prod_{e \in E^c_{w_{\mathbf{i}}}} \mathbb{E}[Z_{1,2}^{N_e^{w_{\mathbf{i}}}}] \prod_{e \in E^s_{w_{\mathbf{i}}}} \mathbb{E}[Y_1^{N_e^{w_{\mathbf{i}}}}] \prod_{e \in E^c_{w_{\mathbf{i'}}}} \mathbb{E}[Z_{1,2}^{N_e^{w_{\mathbf{i'}}}}] \prod_{e \in E^s_{w_{\mathbf{i'}}}} \mathbb{E}[Y_1^{N_e^{w_{\mathbf{i'}}}}]\end{aligned} \tag{3.6}$$

In view of (3.6), it now suffices to show that $\mathscr{W}^{(2)}_{k,t}$ is empty for $t \geq k+2$ since N^{k+2} beats $C_{N,t} \approx N^t$ for $t \geq k+2$. We will show the stronger $t \geq k+1$.
Note that if $a \in \mathscr{W}^{(2)}_{k,t}$, then G_a is a connected graph with t vertices, and at most k edges, which is impossible for $t > k+1$. Considering when $t = k+1$, G_a is a tree and each edge must be visited by a exactly twice. But since the path

from w_1 in tree G_a is closed, we must visit each edge an even number of times. Thus the set of edges from w_1 is disjoint to the set of edges from w_2, creating a contradiction, and we are done.□

4 Final Proof of Wigner's:

To conclude the proof of Wigner's Semicircular Theorem for real Wigner matrices, we must check that for bounded continuous functions f

$$\lim_{N\to\infty} \langle L_N, f\rangle = \langle \sigma, f\rangle \quad \text{in probability}$$

By Chebyshev's Inequality,

$$\begin{aligned}
\mathbb{P}(\langle L_N, |x|^k \mathbf{1}_{\{|x|>B\}}\rangle > \epsilon) &\le \frac{1}{\epsilon}\mathbb{E}[\langle L_N, |x|^k \mathbf{1}_{\{|x|>B\}}\rangle] \\
&= \frac{1}{\epsilon}\langle \overline{L}_N, |x|^k \mathbf{1}_{\{|x|>B\}}\rangle \\
&\le \frac{1}{\epsilon}\langle \overline{L}_N, \frac{x^{2k}}{B^k}\rangle \\
&= \frac{\langle \overline{L}_N, x^{2k}\rangle}{\epsilon B^k}
\end{aligned}$$

Hence by Lemma 1 we have,

$$\begin{aligned}
\limsup_{N\to\infty} \mathbb{P}(\langle L_N, |x|^k \mathbf{1}_{\{|x|>B\}}\rangle > \epsilon) &\le \limsup_{N\to\infty} \frac{\langle \overline{L}_N, x^{2k}\rangle}{\epsilon B^k} \\
&\le \frac{\langle \sigma, x^{2k}\rangle}{\epsilon B^k} = \frac{C_k}{\epsilon B^k} \le \frac{4_k}{\epsilon B^k}
\end{aligned}$$

Now we set $B = 5$ and note the LHS is increasing in k

$$0 \le \limsup_{N\to\infty} \mathbb{P}(\langle L_N, |x|^k \mathbf{1}_{\{|x|>B\}}\rangle > \epsilon) \le \frac{1}{\epsilon}(\frac{4}{5})^k \xrightarrow{k\to\infty} 0$$

Hence $\limsup_{N\to\infty} \mathbb{P}(\langle L_N, |x|^k \mathbf{1}_{\{|x|>B\}}\rangle > \epsilon) = 0 \qquad (4.1)$

Fix an f with support $[-5, 5]$ and $\delta > 0$, then by the Weierstrass Approximation Theorem, we can find a polynomial $\mathcal{Q}_\delta(x) = \sum_{i=0}^{L} c_i x^i$ such that

$$\sup_{x:|x|\le\delta} |\mathcal{Q}_\delta(x) - f(x)| \le \frac{\delta}{8}$$

Then,

$$\begin{aligned}
\mathbb{P}(|\langle L_N, f\rangle - \langle \sigma, f\rangle| > \delta) \le\ & \mathbb{P}(|\langle L_N, \mathcal{Q}_\delta\rangle - \langle \overline{L}_N, \mathcal{Q}_\delta\rangle| > \frac{\delta}{4}) \\
&+ \mathbb{P}(|\langle \overline{L}_N, \mathcal{Q}_\delta\rangle - \langle \sigma, \mathcal{Q}_\delta\rangle| > \frac{\delta}{4}) \\
&+ \mathbb{P}(|\langle L_N, \mathcal{Q}_\delta \mathbf{1}_{\{|x|>B\}}\rangle| > \frac{\delta}{4})
\end{aligned}$$

The first term tends to zero in N by Lemma 2, the second term tends to zero in N by Lemma 1, and the third term tends to 0 in N by (4.1). Hence this finishes the proof. $\square$