Wigner's Semicircle Distribution - Lecture Notes

David Veitch
University of Toronto
david.veitch@mail.utoronto.ca

04/09/2021

Acknowledgement

Partially based on the lecture notes of (Kemp 2016) [2]

1 Motivation - Image Denoising

- one example in (Basu et al. 2010) [1]

- take a picture, and something is corrupting it (e.g. blur)

- would like to remove blur

- let X be a sqaure matrix with entries representing the color of the pixel

- make assumption noise is Gaussian

- look at local neighbourhoods in the image, assume pixels are random (i.e. noise)

- know distribution of eigenvalues of random matrix (Wigner), any eigenvalues outside of this are 'non-random'

Original Picture(L), Noisy Version(M), After Image Processing(R)

Figure 1: Source: `https://math.uni.lu/eml/projects/reports/random-matrices.pdf`

1.1 Other Applications

- study of "wave functions of quantum mechanical systems which are assumed to be so complicated that statistical considerations can be applied to them" (Wigner 1955) [4]

- like the CLT for free probability (studying situations where random variables $XY = YX$ not necessarily true)

2 Important Definitions and Results

Definition 2.1 (Wigner matrix) $\mathbf{Y}_n$ *a matrix with entries* $\{Y_{ij}\}_{1\leq i,j\leq n}$*, where* $Y_{ij} = Y_{ji}$*.* $\{Y_{ij}\}_{1\leq i,j\leq n}$ *independent random variables where* $Y_{ii} \overset{iid}{\sim} F_1$ *and* $Y_{ij} \overset{iid}{\sim} F_2$*.* $0 < \mathrm{E}[Y_{ij}^2] < \infty$*. Then*

$$\mathbf{X}_n = \frac{1}{\sqrt{n}}\mathbf{Y}_n \tag{1}$$

is a Wigner matrix

Recall: since $\mathbf{X}_n$ real and symmetric $\Rightarrow \mathbf{X}_n$ has n real eigenvalues.

Definition 2.2 (Empirical Law of Eigenvalues) $\mathbf{X}_n$ *a Wigner matrix, its empirical law of eigenvalues* $\mu_{\mathbf{X}_n}$ *is the random discrete probability measure*

$$\mu_{\mathbf{X}_n} = \frac{1}{n}\sum_{j=1}^{n}\delta_{\lambda_j(\mathbf{X}_n)} \tag{2}$$

Definition 2.3 (Wigner's semicircle distribution) *The following probability density function is the Wigner semicircle distribution*

$$\sigma_t(x) = \begin{cases} \frac{1}{2\pi t}\sqrt{(4t - x^2)} & x \in [-2t, 2t] \\ 0 & \text{otherwise} \end{cases}. \tag{3}$$

Theorem 2.4 (Wigner's Semicircle Law) $\mathbf{X}_n$ *be a sequence of Wigner matrices with* $\mathrm{E}[Y_{ij}] = 0$ *and* $E[Y_{ij}^2] = t$*. Then for any* $f : \mathbb{R} \to \mathbb{R}$ *that is continuous and bounded, and for each* $\epsilon > 0$

$$\lim_{n\to\infty} P\left(\left|\int f\, d\mu_{\mathbf{X}_n} - \int f\, d\sigma_t\right| > \epsilon\right) = 0. \tag{4}$$

Note: where $\mathbf{X}_n = U_n^T \Lambda_n U_n$, using cyclic property of trace, we could rewrite $\int f\, d\mu_{\mathbf{X}_n}$ as

$$\int f\, d\mu_{\mathbf{X}_n} = \frac{1}{n}\sum_{j=1}^{n} f\left(\lambda_j(\mathbf{X}_n)\right) = \frac{1}{n}\sum_{j=1}^{n} f([\Lambda_n]_{jj}) = \frac{1}{n}\mathrm{Tr}\left(f(\Lambda_n)\right) = \frac{1}{n}\mathrm{Tr}(U_n^T f(\Lambda_n)U) = \frac{1}{n}\mathrm{Tr}\left(f(\mathbf{X}_n)\right) \tag{5}$$

Note: essentially means in limit the empirical law of eigenvalues $\approx$ Wigner semi-circle distribution since can approximate indicator function with continuous bounded f

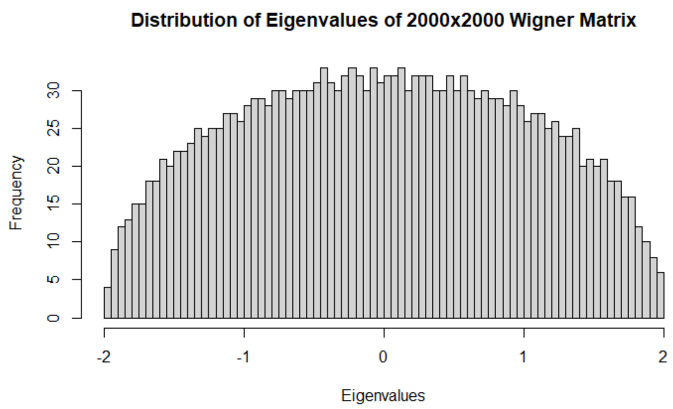

Figure 2: Histogram of eigenvalues for a 2000x2000 simulated Wigner matrix with standard normal entries.

Note: Theorem 2.4 does not really depend on the second moments of diagonals (other than them being finite)

Theorem 2.5 (Wigner's Law for Matrix Moments) $\mathbf{X}_n$ *a sequence of Wigner matrices with* $\mathrm{E}[Y_{ij}] = 0$ *and* $E[Y_{ij}^2] = t \; \forall \; i \neq j$. *Then for fixed* $k \in \mathbb{N}$

$$\lim_{n\to\infty} P\left(\left|\frac{1}{n}\mathrm{Tr}(\mathbf{X}_n^k) - \int x^k \; \sigma_t(x)\right| > \epsilon\right) = 0. \tag{6}$$

Note: Theorem 2.5 is what we will prove, inuitively going to Theorem 2.4 just a matter of approximating f with polynomials.

Fact 2.6 $\int x^k \sigma_t(x) = t^{k/2} \int x^k \sigma_1(x)$ *(t second moment of off-diagonal).*

Fact 2.7 *Let* $m_k = \int x^k \sigma_1(x)$*, by symmetry* $m_{2k+1} = 0 \; \forall \; k,\; m_0 = 1$.

Fact 2.8 $m_{2k} = \frac{2(2k-1)}{k+1} m_{2(k-1)} = C_k = \frac{1}{k+1}\binom{2k}{k}$ *where* C_k *a Catalan number.*

Fact 2.9 C_k *is the number of paths from the bottom left to the upper right of a* $k \times k$ *grid which do not pass above the diagonal [3].*

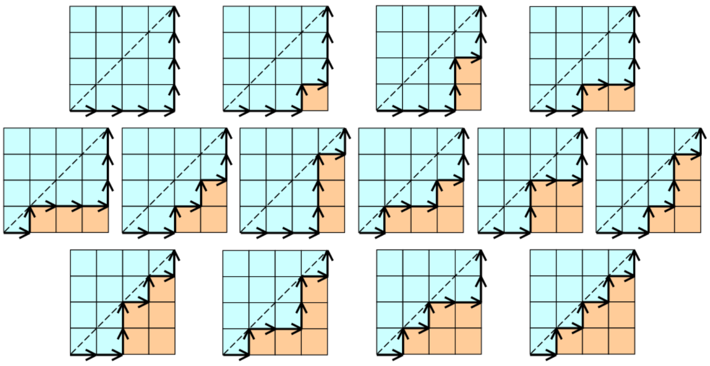

Figure 3: Source: Wikipedia

3 Proof Sketch of Equation 7

To prove Theorem 2.5 we need to show the following

$$\lim_{n\to\infty} \mathrm{E}\left[\mathrm{Tr}\left(\frac{1}{n}\mathbf{X}_n^k\right)\right] = \begin{cases} t^{k/2} C_{k/2} & k \text{ even} \\ 0 & k \text{ odd} \end{cases} \tag{7}$$

$$\lim_{n\to\infty} \mathrm{Var}\left[\mathrm{Tr}\left(\frac{1}{n}\mathbf{X}_n^k\right)\right] = 0 \tag{8}$$

1. Think of matrix multiplication in terms of walks on graphs
2. Show that asymptotically only specific terms enter into the matrix multiplication, and we can take expected values to get the trace

3.1 Expected Value

First note that

$$\mathrm{E}\left[\mathrm{Tr}\left(\frac{1}{n}\mathbf{X}_n^k\right)\right] = \frac{1}{n}\mathrm{E}\left[\mathrm{Tr}\left(\left(\frac{1}{\sqrt{n}}\mathbf{Y}_n\right)^k\right)\right] = \frac{1}{n^{k/2+1}}\mathrm{E}\left[\mathrm{Tr}\left(\mathbf{Y}_n^k\right)\right] \tag{9}$$

Some Notation

- $[n] = \{1, \ldots, n\}$
- $\mathbf{i} = (i_1, \ldots, i_k)$

- $Y_{\mathbf{i}} = Y_{i_1 i_2} Y_{i_2 i_3} \cdots Y_{i_k i_1}$

Definition 3.1 **($G_{\mathbf{i}}, V_{\mathbf{i}}, E_{\mathbf{i}}, w_{\mathbf{i}}$)** *for $\mathbf{i} \in [n]^k$ a k-index $\mathbf{i} = (i_1, \ldots, i_k)$ let the graph $G_{\mathbf{i}}$ have verticies $V_{\mathbf{i}}$ which are distinct elements of $\mathbf{i}$ and let it have edges $E_{\mathbf{i}}$ which are distinct pairs among*

$$\{\{i_1, i_2\}, \{i_2, i_3\}, \ldots, \{i_{k-1}, i_k\}, \{i_k, i_1\}\}.$$

Then let the path $w_{\mathbf{i}}$ be the sequence of edges

$$w_{\mathbf{i}} = (\{i_1, i_2\}, \{i_2, i_3\}, \ldots, \{i_{k-1}, i_k\}, \{i_k, i_1\}\}. \tag{10}$$

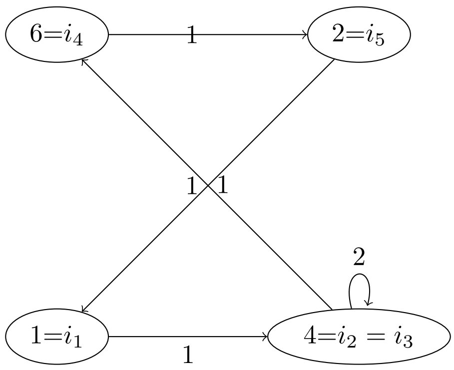

Figure 4: Example of $G_{\mathbf{i}}$ when $\mathbf{i} = \{1, 4, 4, 4, 6, 2\}$

Fact 3.2 (Repeated Matrix Multiplication) *For any $1 \leq i, j \leq n$*

$$[\mathbf{Y}_n^k]_{ij} = \sum_{1 \leq i_2, \ldots, i_k \leq n} Y_{i i_2} Y_{i_2 i_3} \cdots Y_{i_{k-1} i_k} Y_{i_k j} \tag{11}$$

More notation

1. $w_{\mathbf{i}}(e)$ is the number of times each edge is traversed, in Figure 4 $w_{\mathbf{i}}(\{4, 4\}) = 2$, and for all other edges equals 1.
2. $E_{\mathbf{i}}^s$ the set of self edges $E_{\mathbf{i}}^s = \{\{i, i\} \in E_{\mathbf{i}}\}$
3. $E_{\mathbf{i}}^c$ the set of connecting edges $E_{\mathbf{i}}^c = \{\{i, j\} \in E_{\mathbf{i}},\ i \neq j\}$
4. $\mathcal{G}_k$ the set of all pairs (G, w) where G a connected graph with at most k vertices, w the closed walk on G with length k (note that $\mathcal{G}_k$ not defined with respect to specific verticies, i.e. it is independent of n)

All of this notation allows us to write the expected trace in the following way

$$\begin{aligned}
\mathrm{E}\left[\mathrm{Tr}\left(\mathbf{Y}_n^k\right)\right] &= \sum_{i_1=1}^{n} \mathrm{E}\left[[\mathbf{Y}_n^k]_{i_1 i_1}\right] && (12)\\
&= \sum_{1 \leq i_1, i_2, \ldots, i_k \leq n} \mathrm{E}[Y_{i_1 i_2} Y_{i_2 i_3} \cdots Y_{i_k i_1}] && (13)\\
&= \sum_{\mathbf{i} \in [n]^k} \mathrm{E}[Y_{\mathbf{i}}] && (14)\\
&= \sum_{\mathbf{i} \in [n]^k} \prod_{1 \leq i \leq j \leq n} \mathrm{E}\left[Y_{ij}^{w_{\mathbf{i}}(\{i,j\})}\right] && (15)\\
&= \sum_{\mathbf{i} \in [n]^k} \left(\prod_{e_s \in E_{\mathbf{i}}^s} \mathrm{E}\left[Y_{11}^{w_{\mathbf{i}}(e_s)}\right]\right)\left(\prod_{e_c \in E_{\mathbf{i}}^c} \mathrm{E}\left[Y_{12}^{w_{\mathbf{i}}(e_c)}\right]\right) && (16)
\end{aligned}$$

So essentially trace determined by the values of $\mathrm{E}[Y_{\mathbf{i}}]$.

Now since the Y_{ii}, Y_{ij} have fixed distributions, many of the $\mathrm{E}[Y_{\mathbf{i}}]$ are same value which we denote as $\Pi(G,w)$

$$
\begin{aligned}
\mathrm{E}\left[\mathrm{Tr}\left(\mathbf{Y}_n^k\right)\right] &= \sum_{(G,w)\in\mathcal{G}_k} \sum_{\substack{\mathbf{i}\in[n]^k \\ (G_{\mathbf{i}},w_{\mathbf{i}})=(G,w)}} \mathrm{E}\left[Y_{\mathbf{i}}\right] && (17)\\
&= \sum_{(G,w)\in\mathcal{G}_k} \Pi(G,w)\left|\{\mathbf{i}\in[n]^k \mid (G_{\mathbf{i}},w_{\mathbf{i}})=(G,w)\}\right| && (18)
\end{aligned}
$$

where in Equation 18 $|\cdot|$ represents the cardinality of the set.

Returning back to a normalized version of $\mathbf{Y}_n^k$, which is $\mathrm{E}\left[\mathrm{Tr}\left(\frac{1}{n}\mathbf{X}_n^k\right)\right] = \frac{1}{n^{k/2+1}}\mathrm{E}\left[\mathrm{Tr}\left(\mathbf{Y}_n^k\right)\right]$ we get

$$
\frac{1}{n}\mathrm{E}\left[\mathrm{Tr}\left(\mathbf{X}_n^k\right)\right] = \sum_{(G,w)\in\mathcal{G}_k} \Pi(G,w)\frac{\left|\{\mathbf{i}\in[n]^k \mid (G_{\mathbf{i}},w_{\mathbf{i}})=(G,w)\}\right|}{n^{k/2+1}} \tag{19}
$$

Lemma 3.3 (Cardinality of k-indexes) *Given $(G,w)\in\mathcal{G}_k$, with $|G|$ the number of verticies in G then*

$$
\left|\{\mathbf{i}\in[n]^k \mid (G_{\mathbf{i}},w_{\mathbf{i}})=(G,w)\}\right| = n(n-1)\cdots(n-|G|+1) \tag{20}
$$

Remember $\mathrm{E}[Y_{ij}]=0$, therefore only consider w where each edge has been crossed at least twice (denoted $w\geq 2$), so we get

$$
\frac{1}{n}\mathrm{E}\left[\mathrm{Tr}\left(\mathbf{X}_n^k\right)\right] = \sum_{\substack{(G,w)\in\mathcal{G}_k \\ w\geq 2}} \Pi(G,w)\frac{n(n-1)\cdots(n-|G|+1)}{n^{k/2+1}}. \tag{21}
$$

Each edge is crossed at least twice, $|w|=k \Rightarrow$ number of edges used is $\leq k/2 \Rightarrow |G|\leq k/2+1$. And $n(n-1)\cdots(n-|G|+1) = On^{|G|} \Rightarrow$ sequence will be bounded.

Also $|G|$ must be n integer so for an odd k we have $|G|\leq k/2+1/2$. Therefore each term on the RHS of Equation 21 is $O(n^{-1/2})$ and $\mathcal{G}_k$ independent of n therefore

$$
\lim_{n\to\infty}\frac{1}{n}\mathrm{E}\left[\mathrm{Tr}\left(\mathbf{X}_n^k\right)\right]\to 0. \tag{22}
$$

Fact 3.4 ($(G,w)\in\mathcal{G}_k$ with $w\geq 2$)

1. *If there is a self edge in G then $|G|\leq k/2$*
2. *If there is an edge in G with $w(e)\geq 3$ then $|G|\leq k/2$*

This fact means those two cases are asympotically 0! Therefore only care about $\mathcal{G}_k^{k/2+1}$ where G has $k/2+1$ vertices, no self-edges, and walk crosses each edge exactly 2 times. This gives us

$$
\begin{aligned}
\frac{1}{n}\mathrm{E}\left[\mathrm{Tr}\left(\mathbf{X}_n^k\right)\right] &= \left(\sum_{(G,w)\in\mathcal{G}_k^{k/2+1}} \Pi(G,w)\frac{n(n-1)\cdots(n-|G|+1)}{n^{k/2+1}}\right) + O_k(n^{-1}) && (23)\\
\lim_{n\to\infty}\frac{1}{n}\mathrm{E}\left[\mathrm{Tr}\left(\mathbf{X}_n^k\right)\right] &= \sum_{(G,w)\in\mathcal{G}_k^{k/2+1}} \Pi(G,w) && (24)\\
&= \sum_{(G,w)\in\mathcal{G}_k^{k/2+1}} \left(\prod_{e_c\in E^c}\mathrm{E}[Y_{12}^2]\right) && (25)\\
&= t^{k/2}\left|\mathcal{G}_k^{k/2+1}\right| && (26)
\end{aligned}
$$

Can be shown that $\left|\mathcal{G}_k^{k/2+1}\right| = C_{k/2}$ where $C_{k/2}$ a Catalan number.

So we have shown that the expected value of the trace converges to $t^{k/2}C_{k/2}$ which is the k-th moment of the Wigner semicircle distribution. Now we need to prove the variance converges to 0 to get convergence in probability.

3.2 Variance

The variance can be rearranged as

$$\mathrm{Var}\left[\frac{1}{n}\mathrm{Tr}(\mathbf{X}_n^k)\right] = \frac{1}{n^{k+2}} \sum_{\mathbf{i},\mathbf{j}\in[n]^k} (\mathrm{E}[Y_\mathbf{i}Y_\mathbf{j}] - \mathrm{E}[Y_\mathbf{i}]\mathrm{E}[Y_\mathbf{j}]) \tag{27}$$

where the terms in the sum can be viewed as two separate walks on two separate graphs.

Similar to the case of the expectation, we attempt to narrow down the terms of the sum that are non-zero. Denoting $w + w' \geq 2$ where each edge is traversed at least twice, and $\pi(G, w, w')$ as the expected value of this graph and its walks, and M_{2k} as a bound on the maximum expected value over all $\pi(G, w, w')$ (note that there are only finitely many ways to rearrange $2k$ arbitrary vertices so M_{2k} is finite. This gives the bound

$$\mathrm{Var}\left[\frac{1}{n}\mathrm{Tr}(\mathbf{X}_n^k)\right] \leq \sum_{\substack{(G,w,w')\in\mathcal{G}_{k,k} \\ w+w'\geq 2}} \pi(G, w, w')\frac{n^{k+1}}{n^{k+2}} \leq \frac{2M_{2k}}{n}\left|\mathcal{G}_{k,k}\right| \tag{28}$$

Now since $M_{2k}, |\mathcal{G}_{k,k}|$ are independent of n, the bound in Equation 28 is $O_k(n^{-1})$, therefore the variance will converge to 0 as n goes to infinity, proving convergence in probability.

References

[1] G. Basu, K. Ray, and P. K. Panigrahi. Random matrix route to image denoising, 2010.

[2] T. Kemp. Math 247a: Introduction to random matrix theory, Dec 2016. URL `http://www.math.ucsd.edu/~tkemp/247A.Notes.pdf`.

[3] M. Črepinšek and L. Mernik. An efficient representation for solving catalan number related problems. *International Journal of Pure and Applied Mathematics*, 56:589–604, 01 2009.

[4] E. P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Annals of Mathematics*, 62(3):548–564, 1955. ISSN 0003486X. URL `http://www.jstor.org/stable/1970079`.