Introduction to Random Matrix Theory: Global and Local Phenomena

Durham University

Ben Rickard

Supervised by Kohei Suzuki and Mustazee Rahman

Introduction to Random Matrices

Random matrices contain elements randomly sampled from probability distributions. They were first studied by statistician Wishart in the 1920s as sample covariance matrices. Intensive study began in the 1950s with Wigner, motivated by applications in nuclear physics. Physicists Mehta and Dyson also contributed significantly, leading to a formal theory of random matrices in the 1960s. Today, RMT spans across fields like machine learning and analytic number theory. Here we focus on the eigenvalue behaviour of Wigner matrices, a specific class of random matrices.

Wigner Matrices

An $N \times N$ matrix X_N is a Wigner matrix if it has entries

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

where there are two families of independent and identically distributed zero-mean real random variables $\{Z_{i,j}\}_{1 \le i < j}$ and $\{Y_i\}_{1 \le i \le N}$, such that $\mathbb{E}[Z_{1,2}^2] = 1$ and all moments of $Z_{i,j}$ and Y_i are finite.

Eigenvalue Distribution of Wigner Matrices

Since Wigner matrices are symmetric, they have N real, ordered eigenvalues $\{\lambda_i^N\}_{i=1}^N$ such that $\lambda_1^N < \cdots < \lambda_N^N$. The empirical eigenvalue distribution describes the spread of such eigenvalues. We denote this by L_N and it is defined as the probability measure $L_N = \frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i^N}$, where δ is the Dirac distribution. Figure 1 shows the empirical eigenvalue distribution for a Wigner matrix X_{50}. We see the eigenvalues cluster around 0 and become more sparse near the edges ± 2.

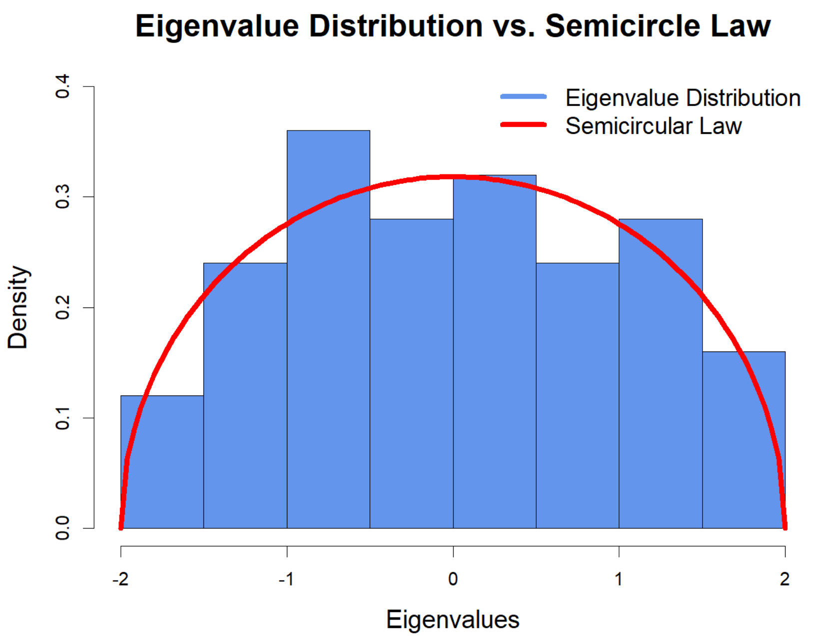

Figure 1. Empirical Eigenvalue Distribution for a 50×50 Wigner Matrix

We question whether the empirical density of all Wigner matrices exhibits a similar semicircle shape, and if a limiting distribution exists as $N \to \infty$. In fact there is such a limiting distribution called the semicircle law, seen in Figure 1, denoted $\sigma(x)dx$ on $\mathbb{R}$ with density

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

Wigner's Semicircle Theorem

For any Wigner matrix X_N, the empirical distribution L_N converges weakly in probability to the semicircle distribution σ i.e. for all $f \in C_b(\mathbb{R})$ and for all $\epsilon > 0$,

$$\lim_{N \to \infty} \mathbb{P}\left(\left| \int f dL_N - \int f d\sigma \right| > \epsilon \right) = 0.$$

Figure 2 shows the closeness of the two distributions for large $N = 5000$ compared to the smaller $N = 50$ in Figure 1. Specifically, we observe the key characteristics of the bulk of eigenvalues clustering near the center, with fewer eigenvalues at the extremes ± 2.

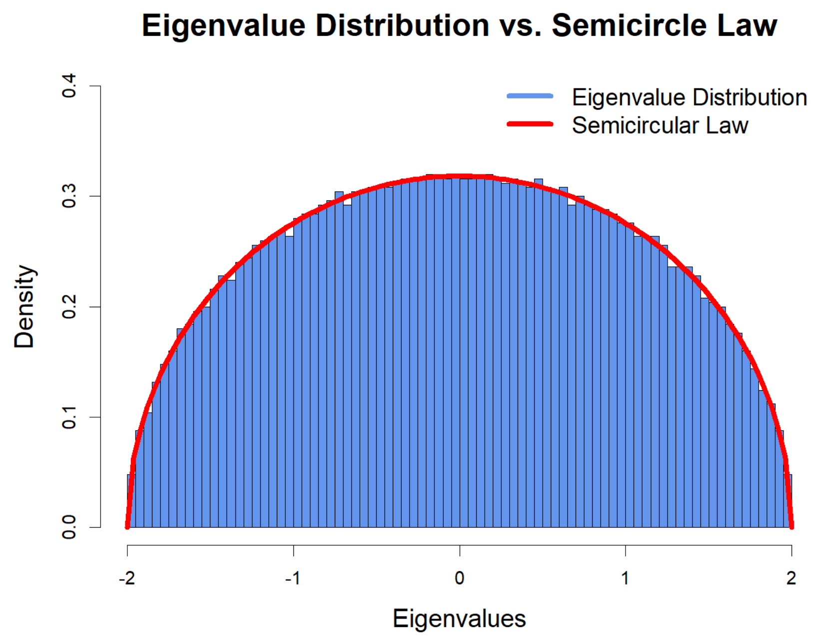

Figure 2. Empirical Eigenvalue Distribution vs. Semicircle Law for N=5000

Global vs. Local Behaviours

So far, our discussion has focused on global phenomena in RMT, where global phenomena refer to properties that describe the entire spectrum of a random matrix or the bulk behaviour of its eigenvalues. These global results provide a macroscopic view of the eigenvalue distribution and are fundamental to understanding the overall structure of random matrices.

However, RMT also considers local phenomena, which focus on microscopic details of the eigenvalue distribution. This includes the analysis of spacings between adjacent eigenvalues, and the distribution of individual eigenvalues. We will now consider the behaviour of the largest eigenvalue of Wigner matrices.

References

[1] G. W. Anderson, A. Guionnet, and O. Zeitouni. *An Introduction to Random Matrices*. Cambridge University Press, 2010.

[2] N. P. Baskerville. *Random Matrix Theory and the Loss Surfaces of Neural Networks*. Phd thesis, University of Bristol, Bristol, 2023. URL `https://arxiv.org/pdf/2306.02108`.

[3] S. Bauman. *The Tracy-Widom Distribution and its Application to Statistical Physics*. MIT Department of Physics, 2017.

Largest Eigenvalue Distribution

In lots of applications of RMT, we are interested in the largest eigenvalue. For example in Principal Component Analysis, a common data science technique, the largest eigenvalue relates to the most important covariate in explaining the variation of a target variable. The following theorem gives the limit distribution of the largest eigenvalue distribution for Wigner matrices.

The Tracy-Widom Theorem

For all $-\infty < t \le \infty$,

$$\lim_{N \to \infty} \mathbb{P}(N^{\frac{2}{3}}(\lambda_N^N - 2) \le t) = \exp\left(- \int_t^\infty (x - t) q(x)^2 dx \right) =: F(t)$$

where:

- q satisfies $q'' = tq + 2q^3$ and also satisfies some asymptotic properties,
- $F(t)$ is called the Tracy-Widom distribution.

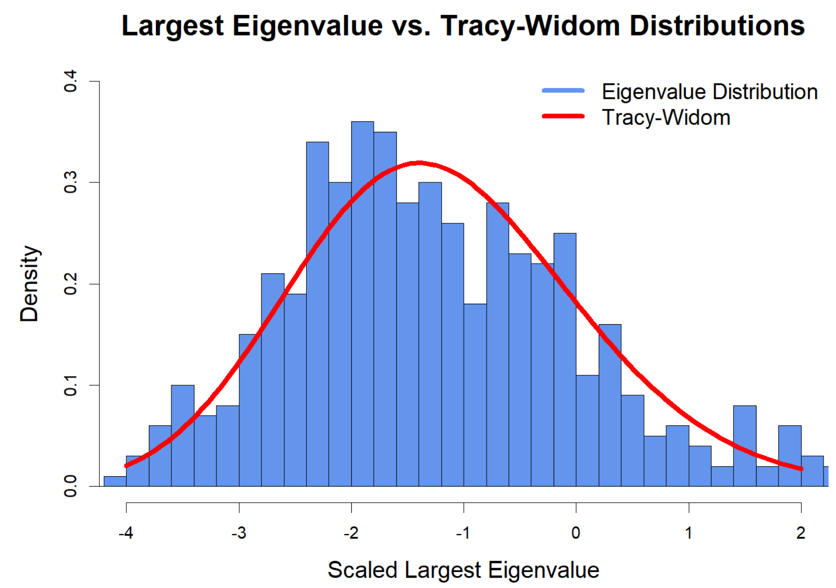

Figure 3. Largest Eigenvalue Distribution vs. Tracy-Widom Distribution for N=1000

Figure 3 depicts the convergence of the empirical largest eigenvalue distribution to the Tracy-Widom distribution for large $N = 1000$. Note that the Tracy-Widom has negative mean, and that it is also asymmetric, with left tail decaying as $e^{-\frac{|x|^3}{12}}$, and the right tail decaying as $e^{-\frac{4}{3}x^{\frac{3}{2}}}$. Comparing to the Gaussian distribution which decays symmetrically at both tails at rate $e^{\frac{x^2}{2}}$, it is clear the Tracy-Widom distribution decays faster than the Gaussian at the left tail, and slower at the right tail.

Conclusion and Further Work

To summarise, we have introduced Wigner matrices and explored both global and local eigenvalue behaviour of them, seeing an example of each. Notably, these results extend beyond Wigner matrices to other matrix ensembles we haven't yet seen. Further directions of study also include considering eigenvalue spacing and other fundamental properties of Wigner matrices. Moreover, as previously mentioned, RMT has a wealth of applications across mathematics, physics, and data science, offering many other avenues for exploration.