13 Edge statistics

In this course, we have mostly studied the spectral densities of matrices in the limit of large N, and typically found that their spectrum has support only in a finite interval $[\lambda_-, \lambda_+]$. The implication of this—that the probability of finding an eigenvalue $\lambda < \lambda_-$ or $\lambda > \lambda_+$ is zero—is only true in the limit of infinite N. For any finite matrix, the probability to find eigenvalues outside the support of the spectral density is small but nonzero, something which can be easily verified by diagonalizing some particular matrix in the computer.

The nature of the edge at finite but large N exhibits many universal features. First, consider large deviations from the edge of the spectrum, where $\lambda_{\text{max}} - \lambda_+$ is order one. Whether λ_{max} is greater or less than the spectral edge results in qualitatively different behavior. The probability that $\lambda_{\text{max}} < w$ for some w is given from the Coulomb gas representation

$$\rho(\boldsymbol{\lambda}) = \frac{1}{Z} e^{-\frac{1}{2}N\sum_{i=1}^N \lambda_i^2 + \frac{1}{2}\sum_{i\neq j}\log|\lambda_i - \lambda_j|} \tag{1}$$

by

$$P(w) = \int_{-\infty}^{w} d\lambda_1 \cdots \int_{-\infty}^{w} d\lambda_N \, \rho(\lambda_1, \ldots, \lambda_N) \tag{2}$$

What is the probability of finding one particular eigenvalue λ_1 far to the outside of the spectrum? Assuming that it is far enough that all the other $N-1$ eigenvalues are weakly affected by its position, we would have

$$\rho(\lambda_1) = \frac{1}{Z}\int d\lambda_2 \, \ldots \, d\lambda_N \, e^{-\frac{1}{2}N\sum_{i=1}^N \lambda_i^2 + \frac{1}{2}\sum_{i\neq j}\log|\lambda_i - \lambda_j|} \tag{3}$$

$$\simeq \frac{1}{Z'} e^{-\frac{1}{2}N\lambda_1^2 + (N-1)\int d\lambda' \, \rho(\lambda')\log|\lambda_1 - \lambda'|} \tag{4}$$

For the GOE ensemble, we can exactly compute the integral over other eigenvalues, which gives

$$\int d\lambda' \, \rho(\lambda')\log|\lambda_1 - \lambda'| = \frac{\lambda}{2}\left(\lambda - \sqrt{\lambda^2 - 2}\right) + \log\left[\frac{1}{2}\left(\lambda + \sqrt{\lambda^2 - 2}\right)\right] - \frac{1}{2} \tag{5}$$

Therefore, summing this integral with the quadratic term and dropping a few constant pieces, we have that

$$\rho(\lambda_1) \propto e^{-N\Phi_+(\lambda_1)} \tag{6}$$

where

$$\Phi_+(\lambda_1) = -\frac{\lambda}{2}\sqrt{\lambda^2 - 2} + \log\left[\frac{1}{2}\left(\lambda + \sqrt{\lambda^2 - 2}\right)\right] \tag{7}$$

This as the form of a large deviation function, dictating that the probability of seeing eigenvalues in this regime shrinks exponentially with N.

The probability of seeing all eigenvalues below some $w < \lambda_+$ also has a large deviation form, though it is more complicated to work out. This can be done from the Coulomb gas picture and the probability $P(w)$ above, where now the Coulomb gas must be considered as confined in a region bounded by a wall at $\lambda = w$. This wall pushes the density of eigenvalues to smaller support, which causes some nonzero fraction of them to be shifted. Since the contributed to the Coulomb gas from each eigenvalue is order N and order N of them are moved by the wall, the large deviation for smaller maximum eigenvalue takes the form

$$P(w) \simeq e^{-N^2\Phi_-(w)} \tag{8}$$

where now the large deviation principle operates at N^2. Therefore, it is much more probable to find eigenvalues outside of the spectrum than it is to see the spectrum compressed.

This accounts for the statistics of large deviations from the spectrum, where the difference between the maximum eigenvalue and the spectral edge is of order one. However, there are also fluctuations that survive at arbitrarily high order in N, but only for deviations $\lambda_{\text{max}} - \lambda_+$ that shrink as N grows. We can argue what this scaling should be in general. Suppose that near the spectral edge, the spectral density vanishes like $\rho(\lambda) \simeq (\lambda_+ - \lambda)^\theta$. For the ensembles we have been studying in the course, $\theta = \frac{1}{2}$. According to the infinite-N spectral density, the probability of seeing an eigenvalue within a small region $\Delta\lambda$ of the spectral edge therefore scales like

$$p(\Delta\lambda) = \int_{\lambda_+-\Delta\lambda}^{\lambda_+} d\lambda\,\rho(\lambda) \propto \Delta\lambda^{\theta+1} \tag{9}$$

However, the sharpness of the infinite-N spectral density should break down for probabilities of order N^{-1}, where the actual discreteness of the eigenvalues will assert itself. Therefore, we expect deviation from the asymptotic spectral density when

$$p(\Delta\lambda) \propto \Delta\lambda^{\theta+1} \sim \frac{1}{N} \qquad \mapsto \qquad \Delta\lambda \sim N^{-\frac{1}{1+\theta}} \tag{10}$$

For $\theta = \frac{1}{2}$, this gives $N^{-\frac{2}{3}}$. We therefore expect to find nontrivial edge distribution when

$$\Phi_{\text{TW}}(u) = \lim_{N\to\infty} N^{\frac{1}{3}}\rho_N(\lambda_+ + uN^{-\frac{2}{3}}) \tag{11}$$

The form of this distribution is called the Tracy–Widom distribution. The distribution is known implicitly for GOE matrices: if $q(s)$ is the solution to the Painlevé II equation

$$q''(s) = 2q(s)^3 + sq(s) \tag{12}$$

with boundary condition $q(s) = \text{Ai}(s)$ as $s \to \infty$ with $\text{Ai}(s)$ the Airy function, then its cumulative distribution is given by

$$\mathcal{F}(x) = \exp\left[-\frac{1}{2}\int_x^\infty ds\,[(s-x)q(s)^2 + q(s)]\right] \tag{13}$$

where $\Phi(u) = \mathcal{F}'(u)$. For the case of GUE matrices, there is an explicit form for Φ when $u > 0$, given by

$$\Phi_{\text{GUE}}(u) = \text{Ai}'(u)^2 - u\,\text{Ai}(u)^2 \tag{14}$$

which has the same asymptotic behavior as the GOE case. The asymptotic distribution Φ has the form

$$\Phi(u) \simeq \begin{cases} e^{-\frac{1}{24}|u|^3} & u \ll 0 \\ e^{-\frac{2}{3}|u|^{\frac{3}{2}}} & u \gg 0 \end{cases} \tag{15}$$

Even in this scaling regime, it is much more probable to find the maximum eigenvalue above the spectral edge than below it, with a steeper power law for negative deviations.