Hard Edge Local Statistics

Connor MacMahon

April 2025

1 Introduction and Global Density

Suppose that in a statistical experiment one takes M i.i.d. samples $\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_M$ with mean $\vec{\mu}$ and fixed $N \times N$ covariance Σ (the null case is $\Sigma = I$). Each of the $\vec{y}_i$ is an $N \times 1$ vector. In an actual experiment, one would take $M > N$. That is, the number of trials exceeds the number of features.

Following [1], consider the sample covariance matrix $\frac{1}{M}XX^T$ where X has columns $\vec{y}_i - \overline{y}$, where $\overline{y}$ is the sample mean. In the null case, the eigenvalues $0 \leq \lambda_1 \leq \lambda_2 \leq \ldots \lambda_N$ are distributed according to the Marchenko-Pastur law:

$$\frac{1}{W_{\alpha\beta N}} \prod_{\ell=1}^{N} \lambda_\ell^{\beta\alpha/2} e^{-\beta\lambda_\ell/2} \prod_{1 \leq j < k \leq N} |\lambda_j - \lambda_k|^\beta.$$

Here $\alpha = M - N + 1 - (2/\beta)$ and $W_{\alpha\beta N}$ a normalizing constant. The parameter β corresponds to real, complex, and quaternionic for $1, 2, 4$ as with the Wigner case. It should be stated that the eigenvalues in this case are singular values, and the large eigenvalues correspond to principal directions in the given dataset, as would be used in a procedure like PCA. Below, however, it is the small eigenvalues which are of interest.

2 The Determinantal Kernel

As in the Wigner case, the joint density above (for $\beta = 2$) arises as $\det(K(\lambda_i, \lambda_j))_{1 \leq i \leq j \leq N}$ for a kernel K of a determinantal point process. Henceforth, set $\beta = 2$. Now, define (as in [2]):

$$\psi_n(x) := \left(\frac{n!}{\Gamma(n + \alpha + 1)}\right)^{1/2} e^{-x/2} x^{\alpha/2} L_n^{(\alpha)}(x),$$

where the $L_n^{(\alpha)}(x)$ is a Laguerre polynomial. Then there is the following Christoffel-Darboux formula ([2]):

$$K_N^{(\alpha)}(x, y) = \sum_{j=0}^{N-1} \psi_j(x)\psi_j(y) = \frac{(N-1)!}{\Gamma(N+\alpha)} \left(\frac{\Gamma(N+\alpha+1)\Gamma(N+\alpha)}{N!(N-1)!}\right)^{1/2} \left(\frac{\psi_{N-1}(x)\psi_N(y) - \psi_N(x)\psi_{N-1}(y)}{x - y}\right).$$

for the relevant kernel K. Since this density incorporates Laguerre polynomials, the above random matrix model is called the Laguerre ensemble. It is of independent interest, but if one sends $\beta \to \infty$ the eigenvalues converge to the roots of Laguerre polynomials.

3 The Hard Edge

As $M/N \to 1$, one obtains a limiting empirical distribution of the eigenvalues as in the Wigner case called the Marchenko-Pastur law. However, observe that since at each stage the eigenvalues are non-negative, this is also true of the limiting law.

While in the Wigner case, eigenvalue density can (with low probability) escape the semicircle for a large Wigner matrix, the Eigenvalues of a Wishart matrix defined above can never go below zero. Hence, 0 is referred to as the hard edge.

What can be said about local statistics at the hard edge? First, one must figure out the correct scaling. Set $R = N + (\alpha + 1)/2$. The following asymptotic may be derived by observing that $\psi_n(x^2)$ solves an analogue of Bessel's equation. Writing this function as a sum of the fundamental solutions (Bessel functions of the first kind J_α) and a particular solution to the non-homogeneous equation and using known asymptotics for Bessel functions ([3]) gives:

$$e^{-x/2}x^{\alpha/2}L_N^\alpha(x) \sim R^{-\alpha/2}\frac{\Gamma(N+\alpha+1)}{N!}J_\alpha(2\sqrt{Rx}) + O(N^{\alpha/2-3/4}).$$

From this, it becomes evident that the scaling which achieves level spacing of x is taking $x/4N$. Using this asymptotic, Stirling's estimate, and the Christoffel-Darboux formula one arrives at the hard edge kernel (as in [2] or [1]):

$$K^{(\alpha)}(x,y) = \lim_{N\to\infty} K_N^{(\alpha)}(\frac{x}{4N}, \frac{y}{4N}) = \frac{J_\alpha(\sqrt{x})\sqrt{y}J_\alpha'(\sqrt{y}) - \sqrt{x}J_\alpha'(\sqrt{x})J_\alpha(\sqrt{y})}{2(x-y)}.$$

This describes the eigenvalues around 0 as a determinantal point process. The associated ℓ-point correlation functions may then be thought of as containing the statistics of the ℓ smallest eigenvalues.

4 The Semiclassical Perspective:

In order to get a handle on this formula, intuitive arguments as in [4] involving differential operators may be made. Observe that for each n, the function ψ_n is an eigenfunction with eigenvalue n for the differential operator:

$$Lf := -xf'' - f' + \left(\frac{x^2 + \alpha^2 - 2x(1+\alpha)}{4x}\right)f.$$

It follows now from the Christoffel-Darboux formula:

$$K_N^{(\alpha)}(x,y) = \sum_{j=0}^{N-1} \psi_j(x)\psi_j(y)$$

that the kernel $K_N^{(\alpha)}$ is the integral kernel of the spectral projection onto the eigenspaces of L corresponding to the first N eigenvalues. I.e. the eigenvalues in the interval $[0, N]$. Equivalently, it is the projection onto the eigenvalues of L/N in the interval $[0, 1]$.

Can the limiting kernel be realized as a spectral projection as well? Perhaps for some "limiting" differential operator? Substitute the scaling $x/4N$ into the differential operator L to obtain:

$$0 \le L = -4Nxf'' - 4Nf' + f\left(\frac{(x/4N)^2 + \alpha^2 - 2(x/4N)(1+\alpha)}{4(x/4N)}\right) \le N,$$

where the inequality denotes the range of the desired spectral projection. Simplifying a bit yields:

$$0 \le -4Nxf'' - 4Nf' + f\left(\frac{x/4N}{4} + \frac{\alpha^2 N}{x} - \frac{1+\alpha}{2}\right) \le N.$$

Now, divide both sides by N to obtain

$$0 \le L = -4xf'' - 4f' + f\left(\frac{x/4N^2}{4} + \frac{\alpha^2}{x} - \frac{1+\alpha}{2N}\right) \le 1.$$

Sending $N \to \infty$ yields the following:

$$0 \le -4xf'' - 4f' + \frac{\alpha^2}{x}f \le 1.$$

Denote the operator in the middle Tf.

The hope is that the limiting kernel above expressed in terms of Bessel functions J_α is the integral kernel of the spectral projection of T onto $[0,1]$.

This is indeed the case. First, observe that the functions $J_\alpha(\pm\sqrt{tx})$ are eigenfunctions for T with corresponding eigenvalue $\pm t$. Hence, the limit operator has continuous spectrum. Moreover, it follows from the orthogonality relations [5] for Bessel functions and a u-substitution that:

$$\int_0^1 J_\alpha(\sqrt{tx})J_\alpha(\sqrt{ty})dt = \frac{J_\alpha(\sqrt{x})\sqrt{y}J'_\alpha(\sqrt{y}) - \sqrt{x}J'_\alpha(\sqrt{x})J_\alpha(\sqrt{y})}{2(x-y)} = K^{(\alpha)}(x,y),$$

the limit kernel! Since the left hand side is an integral and not a sum, however, it is not immediately evident that this is actually the integral kernel of the spectral projection onto $[0,1]$ for T. This is verified as follows via Fubini:

$$\begin{aligned}\int_{-\infty}^{\infty} K^{(\alpha)}(x,y)K^{(\alpha)}(y,z)dy &= \int_0^1\int_0^1\int_{-\infty}^{\infty} J_\alpha(\sqrt{t_1x})J_\alpha(\sqrt{t_1y})J_\alpha(\sqrt{t_2y})J_\alpha(\sqrt{t_2z})dydt_1dt_2\\ &= \int_0^1\int_0^1 J_\alpha(\sqrt{t_1x})J_\alpha(\sqrt{t_2z})\left(\int_{-\infty}^{\infty} J_\alpha(\sqrt{t_1y})J_\alpha(\sqrt{t_2y})dy\right)dt_1dt_2\\ &= \int_0^1\int_0^1 J_\alpha(\sqrt{t_1x})J_\alpha(\sqrt{t_2z})\delta_{t_1=t_2}dt_1dt_2\\ &= \int_0^1 J_\alpha(\sqrt{tx})J_\alpha(\sqrt{tz})dt = K^{(\alpha)}(x,z).\end{aligned}$$

Hence, $K^{(\alpha)}$ is indeed a projection kernel. Now, one needs to be careful as the J_α are not necessarily in L^2, but this is nonrigorous anyway. What is it projecting onto? It projects onto the eigenvalues t which land in $[0,1]$.

5 Bibliography

References

1. Forrester, P. J. *Log-Gases and Random Matrices* (Princeton University Press, 2010).
2. Forrester, P. J. The Spectrum Edge of Random Matrix Ensembles. *Nuclear Physics* (1993).
3. Szegö, G. *Orthogonal Polynomials* (American Mathematical Society, 1939).
4. Tao, T. The Dyson and Airy Kernels of GUE Via Semiclassical Analysis (2010).
5. Dobrushkin, V. Orthogonality of Bessel Functions, Mathematica Tutorial Part 2.7 (2025).