

Lectures on Random Matrices (Spring 2025)

Lecture 5: Determinantal Point Processes and the GUE

Leonid Petrov

Wednesday, February 5, 2025*

Contents

1	Recap	1
2	Discrete determinantal point processes	2
2.1	Definition and basic properties	2
3	Determinantal structure in the GUE	3
3.1	Correlation functions as densities with respect to Lebesgue measure	3
3.2	The GUE eigenvalues as DPP	3
3.2.1	Setup	3
3.2.2	Writing the Vandermonde as a determinant	4
3.2.3	Orthogonalization by linear operations	4
3.2.4	Rewriting the density in determinantal form	5
3.3	Christoffel–Darboux Formula	6
3.4	Double Contour Integral Representation and Steepest Descent	7
4	Summary and Outlook	7
E	Problems (due DATE)	8

1 Recap

In [Lecture 4](#) we discussed global spectral behavior of tridiagonal $G\beta E$ random matrices, and obtained the Wigert semicircle law for the eigenvalue density.

In this lecture we shift our focus to another powerful technique in random matrix theory: the theory of *determinantal point processes* (DPPs). In the $\beta = 2$ (GUE) case the joint eigenvalue distributions can be written in determinantal form. We begin by discussing the discrete version of determinantal processes, and then derive the correlation kernel for the GUE using orthogonal polynomial methods. Finally, we show how the Christoffel–Darboux formula yields a

*[Course webpage](#) • [Live simulations](#) • [TeX Source](#) • Updated at 23:31, Saturday 1st February, 2025

compact representation of the kernel and indicate how one may represent it as a double contour integral—an expression well suited for steepest descent analysis in the large- n limit.

2 Discrete determinantal point processes

2.1 Definition and basic properties

Let \mathfrak{X} be a (finite or countably infinite) discrete set. A *point configuration* on \mathfrak{X} is any subset $X \subset \mathfrak{X}$ (with no repeated points). A random point process is a probability measure on the space of such configurations.

Definition 2.1 (Determinantal Point Process). A random point process P on \mathfrak{X} is called *determinantal* if there exists a function (the *correlation kernel*) $K : \mathfrak{X} \times \mathfrak{X} \rightarrow \mathbb{C}$ such that for any n and every finite collection of distinct points $x_1, \dots, x_n \in \mathfrak{X}$, the joint probability that these points belong to the random configuration is

$$\mathbb{P}\{x_1, \dots, x_n \in X\} = \det \left[K(x_i, x_j) \right]_{i,j=1}^n.$$

Determinantal processes are very useful in probability theory and random matrices. They are a natural extension of Poisson processes, and have some parallel properties. Many properties of determinantal processes can be derived from “linear algebra” (broadly understood) applied to the kernel K . There are a few surveys on them: [Sos00], [HKPV06], [Bor11], [KT12]. Let us just mention two useful properties.

Proposition 2.2 (Gap Probability). *If $I \subset \mathfrak{X}$ is a subset, then*

$$\mathbb{P}\{X \cap I = \emptyset\} = \det \left[I - K_I \right],$$

where K_I is the restriction of the kernel to I . If I is infinite, then the determinant is understood as a Fredholm determinant.

Remark 2.3. The Fredholm determinant might “diverge” (equal to 0 or 1).

Proposition 2.4 (Generating functions). *Let $f : \mathfrak{X} \rightarrow \mathbb{C}$ be a function such that the support of $f - 1$ is finite. Then the generating function of the multiplicative statistics of the determinantal point process is given by*

$$\mathbb{E} \left[\prod_{x \in X} f(x) \right] = \det \left[I + (\Delta_f - I)K \right],$$

where the expectation is over the random point configuration $X \subseteq \mathfrak{X}$, Δ_f denotes the operator of multiplication by f (i.e., $(\Delta_f g)(x) = f(x)g(x)$) and the determinant is interpreted as a Fredholm determinant if \mathfrak{X} is infinite.

Remark 2.5 (Fredholm Determinant — Series Definition). The Fredholm determinant of an operator A on $\ell^2(\mathfrak{X})$ is given by the series

$$\det(I + A) = \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{x_1, \dots, x_n \in \mathfrak{X}} \det \left[A(x_i, x_j) \right]_{i,j=1}^n,$$

where the term corresponding to $n = 0$ is defined to be 1.

3 Determinantal structure in the GUE

3.1 Correlation functions as densities with respect to Lebesgue measure

In the discrete setting discussed above the joint probabilities of finding points in specified subsets of \mathfrak{X} are given by determinants of the kernel evaluated at those points. When the underlying space is continuous (typically a subset of \mathbb{R} or \mathbb{R}^d), one works instead with correlation functions which serve as densities with respect to the Lebesgue measure.

Let $X \subset \mathbb{R}$ be a random point configuration. The n -point correlation function $\rho_n(x_1, \dots, x_n)$ is defined by the relation

$$\begin{aligned} \mathbb{P}\{\text{there is a point in each of the infinitesimal intervals } [x_i, x_i + dx_i], i = 1, \dots, n\} \\ = \rho_n(x_1, \dots, x_n) dx_1 \cdots dx_n. \end{aligned}$$

For a determinantal point process the correlation functions take a determinantal form:

$$\rho_k(x_1, \dots, x_k) = \det \left[K(x_i, x_j) \right]_{i,j=1}^k.$$

Remark 3.1. The reference measure does not necessarily have to be the Lebesgue measure. For example, in the discrete setting, we can also talk about the reference measure, it is the counting measure. The correlation kernel $K(x, y)$ is better understood not as a function of two variables, but as an operator on the Hilbert space $L^2(\mathfrak{X}, d\mu)$, where μ is the reference measure. One can also write $K(x, y) \mu(dy)$ or $K(x, y) \sqrt{\mu(dx) \mu(dy)}$ to emphasize this structure.

This formulation is particularly useful in the continuous setting, as it allows one to express statistical properties of the point process in terms of integrals over the kernel. For example, the expected number of points in a measurable set $A \subset \mathbb{R}$ is given by

$$\mathbb{E}[\#(X \cap A)] = \int_A \rho_1(x) dx,$$

while higher order joint intensities provide information about correlations between points.

3.2 The GUE eigenvalues as DPP

3.2.1 Setup

We start from the joint eigenvalue density for the Gaussian Unitary Ensemble (GUE)

$$p(x_1, \dots, x_n) dx_1 \cdots dx_n = \frac{1}{Z_{n,2}} \prod_{j=1}^n e^{-x_j^2/2} \prod_{1 \leq i < j \leq n} (x_i - x_j)^2 dx_1 \cdots dx_n. \quad (3.1)$$

We will show step by step why this is a determinantal point process,

$$\rho_k(x_1, \dots, x_k) = \det \left[K_n(x_i, x_j) \right]_{i,j=1}^k, \quad k \geq 1,$$

with the kernel defined as

$$K_n(x, y) = \sum_{j=0}^{n-1} \psi_j(x) \psi_j(y),$$

where the functions

$$\psi_j(x) = \frac{1}{\sqrt{h_j}} p_j(x) \sqrt{w(x)}, \quad w(x) = e^{-x^2/2},$$

are constructed from the monic Hermite polynomials $\{p_j(x)\}$ which are orthogonal with respect to the weight $w(x)$:

$$\int_{-\infty}^{\infty} p_j(x) p_k(x) e^{-x^2/2} dx = h_j \delta_{jk}.$$

Recall that “monic” means that the leading coefficient of $p_j(x)$ is 1, and we divide by the norm to make the polynomials orthonormal.

3.2.2 Writing the Vandermonde as a determinant

The product

$$\prod_{1 \leq i < j \leq n} (x_i - x_j)^2$$

is the square of the Vandermonde determinant. Recall that the Vandermonde determinant is given by

$$\Delta(x_1, \dots, x_n) = \prod_{1 \leq i < j \leq n} (x_j - x_i) = \det \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix}.$$

Thus, we have

$$\prod_{1 \leq i < j \leq n} (x_i - x_j)^2 = \left(\det [x_i^{j-1}]_{i,j=1}^n \right)^2.$$

3.2.3 Orthogonalization by linear operations

Since determinants are invariant under elementary row or column operations, we can replace the monomials x^{j-1} by any sequence of monic polynomials of degree $j-1$. In particular, we choose the monic Hermite polynomials $p_{j-1}(x)$ and obtain

$$\det [x_i^{j-1}]_{i,j=1}^n = \det [p_{j-1}(x_i)]_{i,j=1}^n.$$

The orthogonality condition for these polynomials is

$$\int_{-\infty}^{\infty} p_j(x) p_k(x) e^{-x^2/2} dx = h_j \delta_{jk}.$$

We define the functions

$$\phi_j(x) = p_j(x) e^{-x^2/4},$$

and then introduce the orthonormal functions

$$\psi_j(x) = \frac{1}{\sqrt{h_j}} \phi_j(x) = \frac{1}{\sqrt{h_j}} p_j(x) e^{-x^2/4}.$$

Note that here the weight splits as $e^{-x^2/2} = e^{-x^2/4} e^{-x^2/4}$, which is useful in the next step.

3.2.4 Rewriting the density in determinantal form

Substituting the determinant form into the joint density (3.1), we have

$$p(\lambda_1, \dots, \lambda_n) = \frac{1}{Z_{n,2}} \prod_{j=1}^n e^{-\lambda_j^2/2} \left[\det \left[p_{j-1}(\lambda_i) \right]_{i,j=1}^n \right]^2.$$

Incorporate the weight factors into the determinant by writing

$$\prod_{i=1}^n e^{-\lambda_i^2/2} = \prod_{i=1}^n \left(e^{-\lambda_i^2/4} \cdot e^{-\lambda_i^2/4} \right),$$

so that

$$\prod_{i=1}^n e^{-\lambda_i^2/4} \det \left[p_{j-1}(\lambda_i) \right]_{i,j=1}^n = \det \left[\phi_{j-1}(\lambda_i) \right]_{i,j=1}^n.$$

Thus, the joint density becomes

$$p(\lambda_1, \dots, \lambda_n) = \frac{1}{\tilde{Z}_{n,2}} \left[\det \left[\phi_{j-1}(\lambda_i) \right]_{i,j=1}^n \right]^2.$$

This squared-determinant structure is characteristic of determinantal point processes.

A classical result (via the Andréief integration formula) shows that if a joint density is expressed as the square of a determinant, then the marginal k -point correlation function can be written as a determinant with a kernel given by

$$K_n(x, y) = \sum_{j=0}^{n-1} \psi_j(x) \psi_j(y).$$

4. Integrating Out the Other Points and Permutation Considerations

The k -point correlation function is defined by integrating out the remaining $n - k$ variables:

$$\rho_k(x_1, \dots, x_k) = \frac{n!}{(n-k)!} \int_{\mathbb{R}^{n-k}} p(x_1, \dots, x_n) dx_{k+1} \cdots dx_n.$$

Here, the factor $\frac{n!}{(n-k)!}$ accounts for the different ways of choosing k eigenvalues from n . When the determinant is expanded, one encounters sums over permutations:

$$\det \left[\phi_{j-1}(\lambda_i) \right]^2 = \sum_{\sigma, \tau \in S_n} \text{sgn}(\sigma) \text{sgn}(\tau) \prod_{i=1}^n \phi_{\sigma(i)-1}(\lambda_i) \phi_{\tau(i)-1}(\lambda_i).$$

Due to the orthogonality

$$\int_{-\infty}^{\infty} \psi_j(x) \psi_k(x) dx = \delta_{jk},$$

integration over the $n - k$ unobserved eigenvalues forces the indices corresponding to these variables to match (i.e., $\sigma(i) = \tau(i)$ for $i = k + 1, \dots, n$). This pairing effectively reduces the double sum over permutations and results in the kernel

$$K_n(x, y) = \sum_{j=0}^{n-1} \psi_j(x) \psi_j(y).$$

Thus, after integrating out the unobserved eigenvalues and properly accounting for the combinatorial factors and permutation symmetry, the k -point correlation function is given by

$$\rho_k(x_1, \dots, x_k) = \det \left[K_n(x_i, x_j) \right]_{i,j=1}^k.$$

Conclusion

In summary, starting from the joint eigenvalue density (3.1) for the GUE, we performed the following steps:

- (i) Represented the Vandermonde product as a determinant.
- (ii) Replaced the monomials by an orthogonal family of monic Hermite polynomials via linear operations.
- (iii) Incorporated the weight factors to recast the density in a squared-determinant form.
- (iv) Integrated out the unobserved eigenvalues while handling permutation symmetry, leading to the determinantal representation for the k -point correlation functions.

This derivation culminates in the formula:

$$\rho_k(x_1, \dots, x_k) = \det \left[K_n(x_i, x_j) \right]_{i,j=1}^k, \quad K_n(x, y) = \sum_{j=0}^{n-1} \psi_j(x) \psi_j(y),$$

which elegantly encapsulates the correlations among eigenvalues in the GUE.

3.3 Christoffel–Darboux Formula

A major advantage of the determinantal representation is that the sum in (??) can be rewritten in closed form using the Christoffel–Darboux formula. In our context, one obtains

$$K_n(x, y) = \sqrt{w(x)w(y)} \frac{\gamma_{n-1}}{\gamma_n} \frac{p_n(x)p_{n-1}(y) - p_{n-1}(x)p_n(y)}{x - y}, \quad (3.2)$$

where γ_n denotes the leading coefficient of $p_n(x)$ (for monic polynomials, $\gamma_n = 1$ but if one uses an alternate normalization this factor appears).

Remark 3.2. The derivation of the Christoffel–Darboux formula is standard in the theory of orthogonal polynomials; see, e.g., [szego1975orthogonal]. The key idea is that the sum in (??) satisfies a three-term recurrence which then telescopes when writing the difference quotient.

3.4 Double Contour Integral Representation and Steepest Descent

For asymptotic analysis (for example, to derive the sine kernel in the bulk or the Airy kernel at the edge), it is extremely useful to represent the kernel in the form of a double contour integral. One classical route is as follows.

One first expresses the Hermite polynomials via a contour integral representation (see, e.g., the generating function or integral representation for Hermite polynomials):

$$H_n(z) = \frac{n!}{2\pi i} \oint_{\Gamma} e^{2zw-w^2} \frac{dw}{w^{n+1}},$$

with an appropriate choice of contour Γ . Inserting this representation into the Christoffel–Darboux formula (3.2) and interchanging summation and integration (justified by uniform convergence) one obtains a representation of the kernel as

$$K_n(x, y) = \frac{1}{(2\pi i)^2} \int_{C_1} \int_{C_2} \frac{e^{n\Phi(x, \xi) - n\Phi(y, \eta)}}{\xi - \eta} d\xi d\eta, \quad (3.3)$$

where Φ is a certain phase function and the contours C_1, C_2 are chosen so that the integrals converge. (The precise form of Φ depends on the rescaling and normalization.) The representation (3.3) is well suited to a steepest descent (saddle point) analysis in the large- n limit, allowing one to derive universal kernels (such as the sine kernel in the bulk)

$$K_{\text{sine}}(x, y) = \frac{\sin \pi(x - y)}{\pi(x - y)},$$

or the Airy kernel at the spectral edge.

4 Summary and Outlook

In this lecture we:

- Recalled the resolvent (Stieltjes transform) method from Lecture 4 and noted that its analytic completion remains open.
- Introduced discrete determinantal point processes and outlined key properties such as the determinantal form of correlation functions and gap probabilities.
- Derived (via the orthogonal polynomial method) the determinantal kernel for the GUE, first as a finite sum (??) and then using the Christoffel–Darboux formula (3.2) for a more compact representation.
- Indicated how one can represent the kernel as a double contour integral (see (3.3)) and how steepest descent techniques are then used to obtain the universal limiting kernels.

In subsequent lectures we will use these results to study local eigenvalue statistics (the sine and Airy kernels) and discuss further universality aspects of random matrices.

E Problems (due DATE)

References

- [Bor11] A. Borodin, *Determinantal point processes*, Oxford handbook of random matrix theory, 2011. arXiv:0911.1153 [math.PR]. [↑2](#)
- [HKPV06] J.B. Hough, M. Krishnapur, Y. Peres, and B. Virág, *Determinantal processes and independence*, Probability Surveys **3** (2006), 206–229. arXiv:math/0503110 [math.PR]. [↑2](#)
- [KT12] A. Kulesza and B. Taskar, *Determinantal Point Processes for Machine Learning*, Foundations and Trends in Machine Learning **5** (2012), no. 2–3, 123–286. arXiv:1207.6083 [stat.ML]. [↑2](#)
- [Sos00] A. Soshnikov, *Determinantal random point fields*, Russian Mathematical Surveys **55** (2000), no. 5, 923–975. arXiv:math/0002099 [math.PR]. [↑2](#)

L. PETROV, UNIVERSITY OF VIRGINIA, DEPARTMENT OF MATHEMATICS, 141 CABELL DRIVE, KERCHOF HALL, P.O. BOX 400137, CHARLOTTESVILLE, VA 22904, USA
E-mail: lenia.petrov@gmail.com