Homework 3, Fall 2023 18.338

Due 10/18 Wednesday 11:59pm

Please submit your homework via canvas.mit.edu.
If you are submitting .jl or .ipynb files, you must additionally submit .html or .pdf file that captures running notebook or code.

Problem sets

1. Solve 3 out of 16 exercises in the DPP notes.

2. Kernels related to Scaling limits

    (a) (M) The Christoffel-Darboux formula (also see equation (5.6) on page 81) states

    $$\sum_{j=0}^{n} \pi_i(x)\pi_j(y) = \frac{k_n}{k_{n+1}} \frac{\pi_n(x)\pi_{n+1}(y) - \pi_n(y)\pi_{n+1}(x)}{y - x}$$

    where k_n is the leading coefficient of π_n. Let

    $$\pi_j(x) = \frac{H_j(x)}{(2^j \sqrt{\pi} j!)^{1/2}}$$

    for Hermite polynomials H_j, then $k_n/k_{n+1} = \sqrt{n}$. Use the known asymptotics

    $$\lim_{m \to \infty} (-1)^m m^{1/4} \pi_{2m}(x) e^{-x^2/2} = \frac{\cos(\xi)}{\sqrt{\pi}}$$

    $$\lim_{m \to \infty} (-1)^m m^{1/4} \pi_{2m+1}(x) e^{-x^2/2} = \frac{\sin(\xi)}{\sqrt{\pi}}$$

    where $\xi = 2\sqrt{m}x$, to prove

    $$K_{2m}(x, y) = e^{-\frac{1}{2}(x^2+y^2)} \sum_{j=0}^{2m-1} \pi_j(x)\pi_j(y)$$

    converges to the sine kernel, i.e.,

    $$\frac{1}{\sqrt{m}} K_{2m}(x, y) \to \frac{2\sin(\xi_1 - \xi_2)}{\pi(\xi_1 - \xi_2)},$$

    where $\xi_1 = 2\sqrt{m}x, \xi_2 = 2\sqrt{m}y$, as $m \to \infty$.

    (b) (C) Numerically verify the last convergence in Problem 1-(a). It would be enough to compare points $x, y \in [0, 1]^2$.

    (c) (C) Point out the numerical issues on the diagonal and corners. Probably on the diagonal, Christoffel-Darboux needs to be replaced by a derivative approximation. Suggest any alternatives.

    (d) Numerically verify the convergence of K_n to the Airy kernel, i.e.,

    $$\frac{1}{\sqrt{2}n^{1/6}} K_n(\sqrt{2n} + \frac{x}{\sqrt{2}n^{1/6}}, \sqrt{2n} + \frac{y}{\sqrt{2}n^{1/6}}) \to \frac{Ai(x)Ai'(y) - Ai'(x)Ai(y)}{x - y},$$

    where $Ai(x)$ is the Airy function and as $n \to \infty$.

3. (Sampling eigenvalues of the GUE without using the GUE) In this problem we will try to sample eigenvalues of the 5×5 GUE using the Hermite kernel K_n above and the continuous DPP. The kernel K_n will play the role of the K matrix in the DPP.

(a) First sample some 5×5 GUE eigenvalues. (Your normalization should be correct when you have the largest eigenvalue roughly around $\sqrt{2n}$ when you have a large $n \times n$ GUE.) Based on your observation, truncate your eigenvalue space. (We will not care about the space where eigenvalues happen rarely.) A suggestion is $[-5, 5]$, but you can take some larger interval based on your observation.

(b) Then discretize your space into 100 intervals. You can then discretize your Hermite kernel into a 100×100 matrix K. You may need Polynomials.jl and SpecialPolynomials.jl

(c) Take a look at the eigenvalues of K. Numerically compute the eigenvalues of K and describe in what sense K is close to a projection matrix.

(d) The code for DPP sampling algorithm is given in course github homework/DPPsampler.jl. This code works with L matrices. What is a simple modification to the algorithm so that it works with K matrices? (Please do not use $K = L(I + L)^{-1}$.)

(e) Try to sample a DPP from the K you created in (b). (You might want to clamp some numerically zero or one eigenvalues.) What is the size of the sample? Is the sample size constant as the mathematics predicts?

(f) Using the DPP sampling algorithm (you can copy the code from the notes if you want), sample eigenvalues of $1,000,000$ 5×5 GUEs. Histogram them and compare against $1,000,000$ 5×5 GUEs.

(g) Interpret the title statement of this problem, GUE without the GUE, in your own words.

4. (*REQUIRED for course 18 graduate students and those who see themselves as having the soul of a mathematician. WARNING : non trivial but worthwhile and doable*) Narayana numbers, as introduced in equation (4.1) (page 52) of the class note, describes many combinatorial quantities on length $2n$ Dyck paths. Dyck paths of length $2n$ are all the paths consisted of $(1, 1)$ steps and $(1, -1)$ steps, that start at $(0, 0)$ and ends at $(2n, 0)$, that never goes under x-axis. Incredibly, Narayana numbers describe the distribution of (1) number of peaks (2) double ascents (3) ascents in even positions (4) long non-final sequences, in $2n$ Dyck paths. In Kreweras's paper JOINT DISTRIBUTIONS OF THREE DESCRIPTIVE PARAMETERS OF BRIDGES, (where n-bridges are just length $2n$ Dyck paths) these are explained with details. Especially in random matrix context, it is known that moments of Laguerre ensembles are related to the number of ascents in even positions (Chapter 4). Prove that Narayana numbers count the number of ascents in even positions, by doing one of the followings.

   (a) Krewaras, in his paper, wrote that the result is "implied by the set of papers [1], [2] and [3]." Read carefully the three references, and prove that Narayana numbers count the number of ascents in even positions.

   (b) Prove by finding a bijection from the set of all length $2n$ Dyck paths to itself, such that the number of peaks are mapped to the number of even ascents.