Random Matrix Theory, homework 1, due March 3.

Problem 1: the Selberg integral. The purpose of this problem is to calculate the partition function of Gaussian β-ensembles ($\beta \geq 0$), i.e. proving that

$$Z_N^{(\beta)} := \int_{\mathbb{R}^N} |\Delta(\lambda_1,\dots,\lambda_N)|^\beta e^{-N\frac{\beta}{4}\sum_{i=1}^N \lambda_i^2} \mathrm{d}\lambda_1 \dots \mathrm{d}\lambda_N = (2\pi)^{N/2}\left(\frac{\beta N}{2}\right)^{-\frac{N(N-1)\beta}{4}-\frac{N}{2}} \prod_{j=1}^N \frac{\Gamma(1+j\frac{\beta}{2})}{\Gamma\left(1+\frac{\beta}{2}\right)}, \tag{1}$$

where $\Delta(\lambda_1,\dots,\lambda_N) = \prod_{1\leq i<j\leq N}(\lambda_i-\lambda_j)$ is the Vandermonde determinant. Here, $\Gamma(z) = \int e^{-t}t^{z-1}$ for $\mathfrak{Re}(z) > 0$. First, we will prove the Selberg integral formula: for any $\gamma_1, \gamma_2 > -1$ and $\gamma \geq 0$,

$$\begin{aligned} S_N(\gamma_1,\gamma_2,\gamma) &:= \int_{[0,1]^N} \left(\prod_{i=1}^N t_i\right)^{\gamma_1} \left(\prod_{i=1}^N (1-t_i)\right)^{\gamma_2} |\Delta(t_1,\dots,t_N)|^{2\gamma} \mathrm{d}t_1 \dots \mathrm{d}t_N \\ &= \prod_{j=0}^{N-1} \frac{\Gamma(1+\gamma_1+j\gamma)\Gamma(1+\gamma_2+j\gamma)\Gamma(1+(j+1)\gamma)}{\Gamma(2+\gamma_1+\gamma_2+(N+j-1)\gamma)\Gamma(1+\gamma)}. \end{aligned} \tag{2}$$

(i) Prove the Euler integral formula:

$$\int_{[0,1]} t^{\gamma_1}(1-t)^{\gamma_2}\mathrm{d}t = \frac{\Gamma(1+\gamma_1)\Gamma(1+\gamma_2)}{\Gamma(2+\gamma_1+\gamma_2)},$$

by writing $\Gamma(1+\gamma_1)\Gamma(1+\gamma_2)$ as a double integral and making an appropriate change of variables.

(ii) In question (ii) to (vii), assume $\gamma \in \mathbb{N}$. Prove that

$$S_N(\gamma_1,\gamma_2,\gamma) = \sum_{0\leq n_1,\dots,n_N\leq 2\gamma(N-1)} c_{n_1,\dots,n_N} \prod_{j=1}^N \frac{\Gamma(1+\gamma_1+n_j)\Gamma(1+\gamma_2)}{\Gamma(2+\gamma_1+\gamma_2+n_j)}$$

for some coefficients $c_{n_1,\dots,n_N}$ independent of γ_1 and γ_2.

(iii) Prove that if $c_{n_1,\dots,n_N} \neq 0$ then $\sum_{i=1}^N n_i = N(N-1)\gamma$. Assuming additionally that $n_1 \leq \dots \leq n_N$, prove that for any $j \in [\![1,N]\!]$ we have

$$(j-1)\gamma \leq n_j \leq (N+j-2)\gamma.$$

For the first inequality, you can first consider $j = N$ and then observe that $\Delta(t_1,\dots,t_j)$ divides $\Delta(t_1,\dots,t_N)$. For the second inequality, you can write $\Delta(t_1,\dots,t_j)$ in terms of $\Delta(t_1^{-1},\dots,t_j^{-1})$.

(iv) Prove that

$$S_N(\gamma_1,\gamma_2,\gamma) = \frac{P(\gamma_1,\gamma_2)}{Q(\gamma_2)} \prod_{j=0}^{N-1} \frac{\Gamma(1+\gamma_1+j\gamma)\Gamma(1+\gamma_2+j\gamma)\Gamma(1+(j+1)\gamma)}{\Gamma(2+\gamma_1+\gamma_2+(N+j-1)\gamma)\Gamma(1+\gamma)}.$$

where P and Q are polynomials with the same degree in γ_2.

(v) By symmetry in γ_1 and γ_2, prove that P/Q is actually a constant $c(\gamma, N)$.

(vi) By ordering $t_1 \leq \dots \leq t_N$ and conditioning on t_N, prove that

$$S_N(0,0,\gamma) = \frac{1}{\gamma(N-1)+1} S_{N-1}(0,2\gamma,\gamma)$$

(vii) Conclude that (2) holds for any $\gamma \in \mathbb{N}$.

(viii) Prove that (2) holds for any $\gamma > 0$. You can assume the following theorem by Carlson.

If f is analytic on $\mathfrak{Re}(z) \geq 0$, vanishes on $\mathbb{N}$ and $f(z) = \mathrm{O}(e^{\mu z})$ with $\mu < \pi$, then $f = 0$ on $\mathfrak{Re}(z) \geq 0$.

(ix) Prove (1). Hint: $e^{-c\lambda^2} = \lim_{L\to\infty}(1-\lambda/L)^{cL^2}(1+\lambda/L)^{cL^2}$.

Problem 2: large deviations for the largest eigenvalue of Gaussian β-ensembles. In this problem, we consider the ordered eigenvalues of symmetric ($\beta = 1$) or Hermitian ($\beta = 2$) Gaussian matrices, i.e. the probability measure

$$\mathrm{d}\mathbb{P}_N(\lambda_1, \dots, \lambda_N) = \widetilde{Z}_N^{-1} |\Delta(\lambda_1, \dots, \lambda_N)|^\beta e^{-N\frac{\beta}{4}\sum_{i=1}^N \lambda_i^2} \mathbb{1}_{\lambda_1 \leq \dots \leq \lambda_N} \mathrm{d}\lambda_1 \dots \mathrm{d}\lambda_N.$$

Let $\varrho(x) = \frac{1}{2\pi}\sqrt{(4-x^2)_+}$. We want to prove that λ_N satisfies a large deviations principle with good rate function

$$I(x) = \begin{cases} -\beta \int \log|x-y|\varrho(y)\mathrm{d}y + \frac{\beta}{4}x^2 - \frac{\beta}{2} & \text{if } x \geq 2 \\ \infty & \text{if } x < 2 \end{cases}.$$

(i) Give a closed form for $\widetilde{Z}_N$.

(ii) Show that

$$\mathbb{P}_N(\lambda_N > L) \leq \frac{\widetilde{Z}_{N-1}}{\widetilde{Z}_N} \int_L^\infty e^{-N\frac{\beta}{4}\lambda_N^2} \mathrm{d}\lambda_N \int \prod_{i=1}^{N-1} \left(|\lambda_N - \lambda_i|^\beta e^{-\frac{\beta}{4}\lambda_i^2} \right) \mathrm{d}\mathbb{P}_{N-1}(\lambda_1, \dots, \lambda_{N-1}).$$

Using $|x - \lambda_i| e^{-\frac{\lambda_i^2}{4}} \leq C e^{\frac{x^2}{8}}$, conclude that

$$\lim_{L\to\infty} \limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N(\lambda_N > L) = -\infty.$$

(iii) Let $\mathrm{d}\mathbb{P}'_N(\lambda_1, \dots, \lambda_N) = \widetilde{Z'_N}^{-1} |\Delta(\lambda_1, \dots, \lambda_N)|^\beta e^{-(N+1)\frac{\beta}{4}\sum_{i=1}^N \lambda_i^2} \mathbb{1}_{\lambda_1 \leq \dots \leq \lambda_N} \mathrm{d}\lambda_1 \dots \mathrm{d}\lambda_N$. Moreover, denote $\mathscr{B}(\varepsilon)$ the ball of radius ε around ϱ, for the distance $\mathrm{d}(\mu, \nu) = \sup_{\|f\|_{\mathrm{Lip}} \leq 1} |\int f \mathrm{d}\mu - \int f \mathrm{d}\nu|$ (or any distance which metrizes the weak topology). Let $\mathscr{B}_L(\varepsilon)$ denote measures in $\mathscr{B}(\varepsilon)$ supported in $[-L, L]$. Prove that

$$\mathbb{P}_N(x < \lambda_N < L) \leq \mathbb{P}_N(\lambda_1 < -L) + \frac{\widetilde{Z}'_{N-1}}{\widetilde{Z}_N} \left(\int_x^L e^{(N-1)\sup_{\mu \in \mathscr{B}_L(\varepsilon)} (\beta \int \log|\lambda - y| \mathrm{d}\mu(y) - \frac{\beta}{4}\lambda^2)} \mathrm{d}\lambda \right.$$
$$\left. + (L - x) e^{(N-1)\beta \log(2L)} \mathbb{P}'_{N-1} \left(\frac{1}{N-1} \sum_{k=1}^{N-1} \delta_{\lambda_k} \notin \mathscr{B}(\varepsilon) \right) \right)$$

Remember the large deviations principle for the empirical spectral measure. Prove that for any $x > 2$,

$$\limsup_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N(\lambda_N \geq x) \leq -I(x).$$

(iv) Let $x > 2$, and $r, \varepsilon > 0$ such that $2 < r < x - 2\epsilon$. Prove that

$$\mathbb{P}_N(x - \varepsilon < \lambda_N < x + \varepsilon) \geq \frac{\widetilde{Z}'_{N-1}}{\widetilde{Z}_N} \int_{x-\varepsilon}^{x+\varepsilon} \mathrm{d}\lambda_N \int_{[-L,r]^{N-1}} e^{\beta \sum_{k=1}^{N-1} \log|\lambda_N - \lambda_k| - N\frac{\beta}{4}\lambda_N^2} \mathrm{d}\mathbb{P}'_{N-1}(\mathrm{d}\lambda_1, \dots, \lambda_{N-1}).$$

Show that for any $\varepsilon > 0$ and $x > 2$ we have

$$\lim_{\varepsilon \to 0} \liminf_{N\to\infty} \frac{1}{N} \log \mathbb{P}_N(x - \varepsilon < \lambda_N < x + \varepsilon) \geq -I(x).$$

(v) Conclude the proof of the large deviations principle for λ_N.

(vi) Give (with no proof) a large deviations principle for the distribution of $(\lambda_{N-k}, \dots, \lambda_N)$, where $k \geq 1$ is fixed.

Open problem. What would be the speed of a large deviations principle for the empirical spectral measure of a random symmetric Bernoulli matrix? What would be its rate function? Same question at the edge.