PROBLEM SESSION II

LECTURER: VADIM GORIN   TA: EVGENI DIMITROV

July 16, 2018

The following session elaborates Lecture 3. The focus is on deriving the conditional uniform distribution of corners in unitarily invariant measures and the Harish-Chandra-Itzykson-Zuber formula discussed in lecture. The first section below lists several basic facts from linear algebra that are isolated as exercises and will be used in the problems later. Our exposition draws upon [1] and [2] and we refer to these articles for further reading.

1. Preliminaries

For any $N \in \mathbb{N}$ we let $\mathrm{Mat}(\mathbb{C}; N)$ denote the set of $N \times N$ matrices with complex entries. An element $M \in \mathrm{Mat}(\mathbb{C}; N)$ is called *Hermitian* if $M_{ij} = \overline{M}_{ij}$ for $1 \leq i, j \leq N$, i.e. $M = M^*$ (its conjugate transpose). The *Spectral theorem* (see https://en.wikipedia.org/wiki/Spectral_theorem) states that if M is an $N \times N$ Hermitian matrix then

(1) M has an *orthonormal basis* of eigenvectors in $\mathbb{C}^N$;
(2) all eigenvalues of M are *real*, and we denote them $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$;
(3) there is a *unitary* matrix U such that $UMU^* = diag(\lambda_1, \dots, \lambda_N)$ (the diagonal matrix, whose (i,i)-th entry is λ_i.

If $A \in \mathrm{Mat}(\mathbb{C}; N)$ is Hermitian then so is its principal submatrix – the element $B \in \mathrm{Mat}(\mathbb{C}; N-1)$ such that $B_{ij} = A_{ij}$ for $1 \leq i, j \leq N-1$. Let $a_1 \geq a_2 \geq \cdots \geq a_N$ denote the eigenvalues of A and $b_1 \geq b_2 \geq \cdots \geq b_{N-1}$ those of B. From the Spectral theorem there exists a unitary $(N-1) \times (N-1)$ matrix U such that $UBU^* = diag(b_1, \dots, b_{N-1})$. We thus have

$$\begin{bmatrix} U & 0 \\ 0 & 1 \end{bmatrix} \cdot A \cdot \begin{bmatrix} U^* & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} diag(b_1, \dots, b_{N-1}) & Uy \\ y^*U^* & A_{NN} \end{bmatrix}, \tag{1}$$

where y is the $(N-1) \times 1$ column vector $[A_{1N}, \dots, A_{(N-1)N}]^t$.

Exercise 1. Using (1) prove the following formula for the characteristic polynomial of A

$$P_A(z) = \prod_{i=1}^{N}(z - a_i) = \prod_{i=1}^{N-1}(z - b_i) \cdot \left[z - A_{NN} - \sum_{j=1}^{N-1} \frac{\xi_j}{z - b_j} \right], \tag{2}$$

where $\xi_j = |[Uy]_j|^2$ (the square of the absolute value of the j-th entry of Uy) for $j = 1, \dots, N-1$. Use (2) to show that the eigenvalues of A and B satisfy $a_1 \geq b_1 \geq a_2 \geq b_2 \cdots \geq a_{N-1} \geq b_{N-1} \geq a_N$ (written as $a \succeq b$ for short).

Exercise 2. Suppose that V is an $N \times N$ unitary matrix such that $VAV^* = diag(a_1, \dots, a_N)$, and let x be the $N \times 1$ column vector $[V_{N1}, \dots, V_{NN}]^*$. Show that the characteristic polynomial of B satisfies the following formula

$$\prod_{i=1}^{N-1}(z - b_i) = \prod_{i=1}^{N}(z - a_i) \cdot \sum_{i=1}^{N} \frac{|x_i|^2}{z - a_i}. \tag{3}$$

Hint: Conjugate $\begin{bmatrix} z \cdot I_N - A & x \\ x^* & 0 \end{bmatrix}$ with $\begin{bmatrix} V & 0 \\ 0 & 1 \end{bmatrix}$ and compare the resulting determinant with $\det[z \cdot I_{N-1} - B]$.

Exercise 3. Prove the Cauchy determinant formula

$$\det\left[\frac{1}{x_i - y_j}\right]_{i,j=1}^{N} = \frac{\prod_{1\le i<j\le N}(x_j - x_i)(y_i - y_j)}{\prod_{1\le i,j\le N}(x_i - y_j)}. \tag{4}$$

By multiplying both sides of (4) by $(-y_N)$ and letting $y_N \to \infty$ show that

$$\det\left[\frac{1}{x_i - y_j}\Big| 1^N\right] = \frac{\prod_{1\le i<j\le N}(x_j - x_i)\cdot\prod_{1\le i<j\le N-1}(y_i - y_j)}{\prod_{i=1}^{N}\prod_{j=1}^{N-1}(x_i - y_j)}, \tag{5}$$

where the matrix on the left side of (5) is the same as that on the left side of (4) with all entries in the last column replaced by 1's.
Hint: Multiply both sides of (4) by $\prod_{1\le i,j\le N}(x_i - y_j)$ and show that both sides are degree $N(N-1)$ polynomials in (x, y) that are skew-symmetric in x's and y's. Finally, compare the coefficient of $(x_1y_1)^{N-1}(x_2y_2)^{N-2}\cdots(x_{N-1}y_{N-1})^1$ on both sides.

Exercise 4. For $y = (y_1 > y_2 > \cdots > y_N) \in \mathbb{R}^N$ define the *Gelfand Tsetlin polytope* $GT_N(y)$ by

$$GT_N(y) := \{(x^1, \dots, x^N) : x^N = y, x^k = \mathbb{R}^k, x^k \succeq x^{k-1}, 2 \le k \le N\}. \tag{6}$$

The set $GT_N(y)$ can naturally be understood as a bounded convex subset of $\mathbb{R}^{N(N-1)/2}$. Prove that its volume is given by

$$|GT_N(y)| = \prod_{1\le i<j\le N} \frac{y_i - y_j}{j - i}. \tag{7}$$

Hint: Let N_L be the number of integer lattice points in $\mathbb{Z} \times \mathbb{Z}^2 \times \cdots \times \mathbb{Z}^{N-1} \cap GT_N(\lambda^L)$, where $\lambda_i^L = \lfloor L \cdot y_i \rfloor$. Argue that $|GT_N(y)| = \lim_{L\to\infty} L^{-N(N-1)/2} N_L$. Finally, use that $N_L = s_{\lambda^L}(1^N)$ and apply Problem 1 from Session I.

Exercise 5. Let $z = (z_1, \dots, z_N), y = (y_1, \dots, y_N) \in \mathbb{R}^N$ be given. Show that

$$e^{z_N \cdot \sum_{i=1}^N y_i} \cdot \det\left[e^{(z_i - z_N)y_j} - e^{(z_i - z_N)y_{j+1}}\right]_{i,j=1}^{N-1} = \det\left[e^{z_i y_j}\right]_{i,j=1}^{N}. \tag{8}$$

Hint: Divide the j-th column of the matrix on the RHS by $e^{z_N y_j}$ for $j = 1, \dots, N$ and then subtract the $j+1$-th column from the j-th one for $j = 1, \dots, N-1$.

2. Problems

Problem 1. Let $y = (y_1 > y_2 > \cdots > y_N) \in \mathbb{R}^N$ be given. Consider the random matrix $H = U diag(y_1, \dots, y_N)U^*$, where U is a random Haar distributed $N \times N$ unitary matrix. Denote by X_i^j for $1 \le j \le N-1$ and $1 \le i \le j$ the ordered (random) eigenvalues of the principal submatrices of the matrix H. In particular, we assume that $X_1^j \ge X_2^j \ge \cdots \ge X_j^j$ are the ordered eigenvalues of the $j \times j$-th submatrix.

From Exercise 1 we know that $X^1 \preceq X^2 \preceq X^3 \preceq \cdots \preceq X^{N-1} \preceq y$. The random $N(N-1)/2$ dimentional vector $(X^1, \dots, X^{N-1})$ can be understood as a random element in Gelfand Tsetlin polytope $GT_N(y)$. Show that $(X^1, \dots, X^{N-1})$ is uniformly distributed on $GT_N(y)$ by completing the following outline.

1. Use Exercise 2 to show that the distribution of X^{N-1} is the same as the distribution of the ordered roots of the polynomial $\prod_{i=1}^N (z-y_i)\cdot\sum_{i=1}^N \frac{|v_i|^2}{z-y_i}$, where $v=(v_1,\dots,v_N)$ is uniformly distributed on the unit sphere S in $\mathbb{C}^N$ (i.e. $S=\{x\in\mathbb{C}^N : x^*x=1\}$).
2. Let $D_1=\{x\in\mathbb{C}^N : x^*x\le 1\}$ and $\Delta_1=\{w=(w_1,\dots,w_N)\in\mathbb{R}^N : w_i\ge 0, w_1+\cdots+w_N\le 1\}$. Consider the map $f : D_1\to\Delta_1$, given by $f(x_1,\dots,x_N)=(|x_1|^2,\dots,|x_N|^2)$ and show that if A is uniformly distributed in D_1, $f(A)$ is uniformly distributed in Δ_1. Conclude that the distribution of X^{N-1} is the same as the distribution of the ordered roots of the polynomial $\prod_{i=1}^N (z-y_i)\cdot\sum_{i=1}^N \frac{W_i}{z-y_i}$, where $W=(W_1,\dots,W_N)$ is uniformly distributed on $\Sigma=f(S)=\{w=(w_1,\dots,w_N)\in\mathbb{R}^N : w_i\ge 0, w_1+\cdots+w_N=1\}$.
3. We obtain the following system of equations relating the variables W_i and X_i^{N-1} (written X_i for short)

$$\begin{aligned}\frac{W_1}{y_1-X_i}+\cdots+\frac{W_N}{y_N-X_i} &=0 \text{ for } i=1,\dots,N-1\\ W_1+W_2+\cdots+W_N &=1.\end{aligned}\tag{9}$$

Using Cramer's rule and Exercise 3 show that

$$W_k=h_k^y(X_1,\dots,X_{N-1}):=\frac{\prod_{i=1}^{N-1}(y_k-X_i)}{\prod_{i=1,i\neq k}^N (y_k-y_i)} \text{ for } k=1,\dots,N.\tag{10}$$

4. Use the previous part to show that the map $h^y=(h_1^y,\dots,h_{N-1}^y)$ defines a bijective diffeomorphism between $I=(y_2,y_1)\times(y_3,y_2)\times\cdots\times(y_N,y_{N-1})$ and the unit simplex $\Delta=\{w=(w_1,\dots,w_{N-1})\in\mathbb{R}^{N-1} : w_i>0, w_1+\cdots+w_{N-1}<1\}$. Conclude that for some constant $c>0$ (you do not need to compute it yet) you have

$$\mathbb{P}(X\in A)=\int_{A\cap I} c\cdot|\det J(x)|dx_1\cdots dx_{N-1},$$

where J is the Jacobian of the map h^y.
5. Show that

$$J(x)=\left[\frac{\prod_{i=1,i\neq j}^{N-1}(y_k-x_i)}{\prod_{i=1,i\neq k}^N (y_k-y_i)}\right]_{k,j=1}^{N-1}$$

and using Exercise 3 prove that

$$|\det J(x)|=\frac{\prod_{1\le i<j\le N-1}(x_i-x_j)}{\prod_{1\le i<j\le N}(y_i-y_j)}.$$

6. Use Exercise 4 to show that the constant c from part 4 is equal to $(N-1)!$.
7. Prove the statement of the problem by induction on N.
Hint: When going from N to $N+1$ use the fact that conditional on X^N the distribution of $(X^1,\dots,X^{N-1})$ is uniform on $GT_N(X^N)$.

Problem 2. Let $y=(y_1,y_2,\dots,y_N), z=(z_1,z_2,\dots,z_N)\in\mathbb{R}^N$ be given. Consider two deterministic Hermitian matrices Z,Y such that $\operatorname{spec}(Z)=z$ and $\operatorname{spec}(Y)=y$ (i.e. the eigenvalues of Z and Y are $z_1,\dots,z_N$ and $y_1,\dots,y_N$ respectively) and the random matrix $H=UYU^*$, where U is a random Haar distributed $N\times N$ unitary matrix. Prove the Harish-Chandra-Itzykson-Zuber formula

$$\mathbb{E}^{z,y}\left[\exp(\operatorname{tr} Z\cdot H)\right]=\prod_{i=1}^{N-1} i!\cdot\frac{\det[\exp(y_iz_j)]_{i,j=1}^N}{\prod_{1\le i<j\le N}(y_i-y_j)(z_i-z_j)}=: B_y(z_1,\dots,z_N).\tag{11}$$

First approach: Reduce the statement to when Z and Y are diagonal and z's and y's are ordered. Proceed by induction on N. When going from $N-1$ to N use the tower property to write

$$\mathbb{E}^{z,y}\left[\exp(\text{trace } Z\cdot H)\right] = \mathbb{E}^{z,y}\left[\exp(z_N\cdot[\text{tr}Y - \text{tr}\tilde{H}])\cdot\mathbb{E}^{z,y}\left[\exp(\text{tr } \tilde{Z}\cdot\tilde{H})|\text{spec}(\tilde{H}) = (\mu_1,\dots,\mu_{N-1})\right]\right],$$

where $\tilde{H}$ is the $(N-1)\times(N-1)$ principal submatrix of H and $\tilde{Z} = diag(z_1,\dots,z_{N-1})$.

Apply the induction hypothesis to deduce that

$$\mathbb{E}^{z,y}\left[\exp(\text{trace } Z\cdot H)\right] = \mathbb{E}^{z,y}\left[\exp(z_N\cdot[\text{tr}Y - \text{tr}\tilde{H}])\cdot\prod_{i=1}^{N-2} i!\cdot\frac{\det[\exp(\mu_i z_j)]_{i,j=1}^{N-1}}{\prod_{1\le i<j\le N-1}(\mu_i-\mu_j)(z_i-z_j)}\right].$$

Finally, apply Step 4 in Problem 1 to get

$$\mathbb{E}^{z,y}\left[\exp(\text{trace } Z\cdot H)\right] = \int_{y_2}^{y_1}\cdots\int_{y_N}^{y_{N-1}} d\mu_1\cdots d\mu_{N-1}\exp\left(z_N\cdot\left[\sum_{i=1}^N y_i - \sum_{i=1}^{N-1}\mu_i\right]\right)\cdot\prod_{i=1}^{N-1} i!\cdot$$
$$\frac{\det[\exp(\mu_i z_j)]_{i,j=1}^{N-1}}{\prod_{1\le i<j\le N-1}(z_i-z_j)\cdot\prod_{1\le i<j\le N}(y_i-y_j)},$$

change the order of the determinant and integral, perform the integral and use Exercise 5 to finish.

Second approach: As discussed in Lecture 3 we have essentially by definition of s_λ that

$$\prod_{i=1}^{N-1} i!\cdot\frac{\det[\exp(y_i z_j)]_{i,j=1}^N}{\prod_{1\le i<j\le N}(y_i-y_j)(z_i-z_j)} = \lim_{\epsilon\to 0}\frac{s_{\lambda(\epsilon)}(e^{\epsilon z_1},\dots,e^{\epsilon z_N})}{s_{\lambda(\epsilon)}(1^N)},$$

where $\lambda(\epsilon) = (\lfloor y_1\epsilon^{-1}\rfloor, \lfloor y_2\epsilon^{-1},\cdots,\lfloor y_N\epsilon^{-1}\rfloor)$. On the other hand, by the branching relations for Schur polynomials we have

$$\frac{s_{\lambda(\epsilon)}(e^{\epsilon z_1},\dots,e^{\epsilon z_N})}{s_{\lambda(\epsilon)}(1^N)} = \frac{\prod_{i=1}^{N-1} i!\epsilon^{-N(N-1)/2}}{\prod_{1\le i<j\le N}(\lambda_i(\epsilon)-\lambda_j(\epsilon)+j-i)}\sum_{\varnothing\preceq\mu^1\preceq\mu^2\cdots\preceq\mu^{N-1}\preceq\lambda}\epsilon^{N(N-1)/2}\prod_{i=1}^N e^{\epsilon z_i(|\mu^i|-|\mu^{i-1}|)},$$

where $\mu^0 = \varnothing$ is the unique signature of length 0. From Exercise 4 conclude that the factor in front of the sum behaves like $\frac{1}{|GT_N(y)|}$ as $\epsilon\to 0$, and so the whole expression can be viewed as a Riemann sum that asymptotically approximates

$$\frac{1}{|GT_N(y)|}\cdot\int_{GT_N(y)} dx^1\cdots dx^{N-1}\exp\left(\sum_{i=1}^N z_i\cdot\left[|x^i|-|x^{i-1}|\right]\right),$$

where $x^N = y$ and $|x| = x_1+\cdots+x_m$ for an m-dimensional vector. Use Problem 1 to deduce that the above expression is precisely $\mathbb{E}^{z,y}\left[\exp(\text{trace } Z\cdot H)\right]$.

3. Optional problems

The following section contains a couple of problems that are related to the material covered in lectures and are meant for interested readers.

Problem A. Let X be an $N\times N$ matrix of i.i.d. standard complex Gaussians (i.e. $X_{mn} = Z_{mn} + i\tilde{Z}_{mn}$ where $Z_{mn},\tilde{Z}_{mn}$ are normal random variables with mean 0 and variance 1). Define $M = \frac{X+X*}{2}$ – this is a random matrix from the Gaussial Unitary Ensemble (GUE). Denote by X_i^j for $1\le j\le N-1$ and $1\le i\le j$ the ordered (random) eigenvalues of the principal submatrices of the matrix M. In particular, we assume that $X_1^j \geq X_2^j \geq \cdots \geq X_j^j$ are the ordered eigenvalues of the $j \times j$-th submatrix. Prove that the density of $(X^1, \dots, X^N)$ is given by

$$f(x^1, \dots, x^N) = \frac{\mathbf{1}(x^1 \preceq x^2 \preceq \cdots \preceq x^N)}{(\sqrt{2\pi})^N} \cdot \prod_{1 \leq i < j \leq N} (x_i^N - x_j^N) \cdot \prod_{i=1}^N e^{-(x_i^N)^2/2}, \tag{12}$$

where $1(\cdot)$ is 1 if the condition in the brackets is satisfied and 0 otherwise. Using Exercise 4 deduce that the density of X^N is given by

$$f(x^N) = \frac{\mathbf{1}(x^1 \preceq x^2 \preceq \cdots \preceq x^N)}{(\sqrt{2\pi})^N \prod_{i=1}^{N-1} i!} \prod_{1 \leq i < j \leq N} (x_i^N - x_j^N)^2 \cdot \prod_{i=1}^N e^{-(x_i^N)^2/2}. \tag{13}$$

Hint: Proceed by induction on N and for the induction step use the following result

$$\begin{aligned} \mathbb{P}(X^{N+1} \in y^{N+1} + dy | X^1, \dots, X^N) = &\frac{1}{\sqrt{2\pi}} \frac{\prod_{1 \leq i < j \leq N+1} (y_i^{N+1} - y_j^{N+1})}{\prod_{1 \leq i < j \leq N} (X_i^N - X_j^N)} \cdot 1(X^N \prec y^{N+1}) \times \\ &\exp\left(- \sum_{i=1}^{N+1} (y_i^{N+1})^2/2 + \sum_{i=1}^N (X_i^N)^2/2 \right) dy. \end{aligned} \tag{14}$$

The main difficulty is in establishing (14). Use Exercise 1 to deduce that the conditional distribution of X^{N+1} given $X^1, \dots, X^N$ is the same as the distribution of the roots of

$$\prod_{i=1}^N (z - X_i^N) \cdot \left[z - \eta - \sum_{j=1}^N \frac{\xi_j}{z - X_i^N} \right],$$

where $\xi_j = |X_{jN}|^2$ and $\eta = X_{(N+1)(N+1)}$. Note ξ_j are exponentially distributed with parameter 1 and η is normal with mean 0 and variance 1 – all variables are independent.

Conclude that we have the following system of equations

$$\begin{aligned} \xi_k &= -\frac{\prod_{i=1}^{N+1} (X_k^N - X_i^{N+1})}{\prod_{i=1, i \neq k}^N (X_k^N - X_i^N)} \text{ for } k = 1, \dots, N \\ \eta &= \sum_{i=1}^{N+1} X_i^{N+1} - \sum_{i=1}^N X_i^N \end{aligned} \tag{15}$$

and note that the map $h : (\xi_1, \dots, \xi_N, \eta) \to (X_1^{N+1}, \dots, X_{N+1}^{N+1})$ defines a bijective diffeomorphism between $\mathbb{R}_{>0}^N \times \mathbb{R}$ and $I = (-\infty, X_N^N) \times (X_N^N, X_{N-1}^N) \times \cdots (X_2^N, X_1^N) \times (X_1^N, \infty)$. The latter implies

$$\mathbb{P}(X^{N+1} \in A | X^N, \dots, X^1) = \int_{A \cap I} |\det J(x)| \frac{\exp\left[-h_1^{-1}(x) - \cdots - h_N^{-1}(x) - h_{N+1}^{-1}(x)^2/2 \right]}{\sqrt{2\pi}} dx_1 \cdots dx_{N+1},$$

where J is the Jacobian of the map h. Similarly to Step 5 in Problem 1 one computes

$$\begin{aligned} |\det J(x)| &= \left| \prod_{k=1}^N \frac{\prod_{i=1}^{N+1} (X_k^N - x_i)}{\prod_{i=1, i \neq k}^N (X_k^N - X_i^N)} \cdot \frac{\prod_{1 \leq i < j \leq N} (X_i^N - X_j^N) \cdot \prod_{1 \leq i < j \leq N+1} (x_i - x_j)}{\prod_{i=1}^N \prod_{j=1}^{N+1} (X_i^N - x_j)} \right| \\ &= \frac{\prod_{1 \leq i < j \leq N+1} (x_i - x_j)}{\prod_{1 \leq i < j \leq N} (X_i^N - X_j^N)}. \end{aligned}$$

A final observation using (1) shows that

$$h_1^{-1}(x) + \cdots + h_N^{-1}(x) + h_{N+1}^{-1}(x)^2/2 = \sum_{i=1}^{N+1} x_i^2/2 - \sum_{i=1}^N (X_i^N)^2/2,$$

which proves (14).

Problem B. Consider the law of $(\ell_1, \dots, \ell_N)$ as in Problem 5 of Session I. Let L be a large parameter and set $A = aL$, $B = bL$, $C = cL$ while N is kept fixed. Show directly that for an appropriate choice of constants d_1 and d_2 (depending on a, b and c) the distribution of the N-dimensional vector

$$\left\{ \frac{\ell_i - d_1 L}{d_2 \sqrt{L}} \right\}_{i=1}^{N} \tag{16}$$

converges to the density in (12).

Hint: The rescaling constants are $d_1 = \frac{ac}{b+c}$ and $d_2 = \sqrt{\frac{abc(a+b+c)}{(b+c)^3}}$.

Remark: In Lectures 2 and 3 the convergence in Problem B was shown using the method of Schur generating functions (SGF). While the approach in Problem B is more direct it fails to generalize to situations where the SGF is applicable. Both methods show that the distribution of random uniform tilings on the hexagon, restricted to a single line near a *turning point* converges to the eigenvalue distribution of the GUE ensemble. The full convergence to the GUE corners is obtained by combining this result with the fact that conditional on a single line one has uniformity in both the tiling model (essentially by definition) and in the GUE corners (using Problem 1 and the unitary invariance of the ensemble). A good place to see ideas of how this is done is [3].

References

[1] Baryshnikov, Y. "Gues and queues". Probab. Theory Relat. Fields, 119:256-274, 2001.

[2] Neretin, Yu. A. "Rayleigh triangles and non-matrix interpolation of matrix beta-integrals", Sbornik: Mathematics, 194(4):515-540, 2003.

[3] Gorin, V. "From alternating sign matrices to the Gaussian Unitary Ensemble", Comm. Math. Phys., 332(1):437-447, 2014.