UNIVERSITY OF CAMBRIDGE

MATHEMATICAL TRIPOS Part III

Monday, 4 June, 2012 9:00 am to 11:00 am

PAPER 31

RANDOM MATRICES

Attempt no more than ***TWO*** *questions.*

There are ***THREE*** *questions in total.*

The questions carry equal weight.

STATIONERY REQUIREMENTS
Cover sheet
Treasury Tag
Script paper

SPECIAL REQUIREMENTS
None

You may not start to read the questions printed on the subsequent pages until instructed to do so by the Invigilator.

1

Let X be an n-by-m matrix of complex Gaussian variables (we assume $n \geqslant m$), and let $A = X'X$. Then A is a random m-by-m matrix and it is known that the distribution of its eigenvalues is given by the following formula:

$$f_m\left(x_1, \ldots, x_m\right) = c_m \prod_{1 \leqslant j < k \leqslant m} \left(x_j - x_k\right)^2 \prod_{i=1}^{m} x_i^a e^{-x_i}, \;\; x_i \geqslant 0,$$

where $a = n - m$.

(i) Prove the Vandermonde identity:

$$\det\left(x_i^{j-1}\right)_{1 \leqslant i,j \leqslant m} := \prod_{1 \leqslant k < l \leqslant m} \left(x_l - x_k\right).$$

(ii) Let $\widetilde{P_j}(x)$ be arbitrary monic polynomials of degree $j-1$. Explain why it is true that

$$\prod_{1 \leqslant k < l \leqslant m} \left(x_l - x_k\right)^2 = \left(\det\left[\widetilde{P_j}(x_i)\right]_{1 \leqslant i,j \leqslant m}\right)^2,$$

and write down

$$\prod_{1 \leqslant j < k \leqslant m} \left(x_j - x_k\right)^2 \prod_{i=1}^{m} x_i^a e^{-x_i},$$

as the square of a determinant.

(iii) Derive the expression

$$f_m\left(x_1, \ldots, x_m\right) = \widehat{c}_m \det\left[K_m\left(x_i, x_j\right)\right]_{1 \leqslant i,j \leqslant m},$$

for the density of eigenvalues, obtaining an explicit formula for $K_m(x,y)$ in terms of the orthogonal polynomials with respect to the weight $x^a e^{-x}$.

(iv) In order to calculate the marginal densities of eigenvalues, it is necessary to integrate the determinant from part (iii) over some of its variables. Prove the following result.

Theorem. *Suppose that the function $K\left(x,y\right)$ has the following two properties:*

$$\int_{\mathbb{R}} K(x,y)K(y,z)dy = K(x,z),$$

and

$$\int_{\mathbb{R}} K(x,x)dx = r.$$

Then,

$$\int_{\mathbb{R}} \det\left[K\left(x_i, x_j\right)\right]_{1 \leqslant i,j \leqslant n} dx_n = \left(r - n + 1\right) \det\left[K\left(x_i, x_j\right)\right]_{1 \leqslant i,j \leqslant n-1}.$$

(v) Define the correlation functions $R_k(x_1, \ldots, x_k)$ and write down an expression for the correlation function R_k as a determinant of the kernel $K_m\left(x,y\right)$.

2

(i) Write the definition of a Lipschitz function $f : \mathbb{R} \to \mathbb{R}$ with constant L.

(ii) Let X_N be a symmetric random matrix, such that $Y_{ij} = \sqrt{N} X_{ij}$ are i.i.d. for $i \leqslant j$, $\mathbb{E}Y_{ij} = 0$ and $\mathbb{E}Y_{ij}^2 = 1$. Let

$$\widehat{X}_{ij} = X_{ij} 1_{\sqrt{N}|X_{ij}|<C} - \mathbb{E}(X_{ij} 1_{\sqrt{N}|X_{ij}|<C}).$$

Let L_N and $\widehat{L}_N$ denote the empirical probability measure of eigenvalues of X_N and $\widehat{X}_N$, respectively. Assume that f is a Lipschitz function $\mathbb{R} \to \mathbb{R}$ with Lipschitz constant 1. Let λ_i and $\widehat{\lambda}_i$ be ordered eigenvalues of X_N and $\widehat{X}_N$, respectively. Show that

$$\left| < L_N, f > - < \widehat{L}_N, f > \right| \leqslant \frac{1}{N} \sum_{i=1}^{N} |\lambda_i - \widehat{\lambda}_i|.$$

(iii) The Hoffman-Wielandt inequality is a useful tool to estimate $\sum_{i=1}^{N} |\lambda_i - \widehat{\lambda}_i|^2$. State the Hoffman-Wielandt inequality.

(iv) Show that

$$\left| < L_N, f > - < \widehat{L}_N, f > \right| \leqslant \left(\frac{1}{N} \mathrm{Tr} \left(X_N - \widehat{X}_N \right)^2 \right)^{1/2}$$

(v) Show that for every $\varepsilon > 0$, it is possible to find C that depends on ε and the law of Y_{ij} only, such that

$$\mathbb{P}\{ | < L_N, f > - < \widehat{L}_N, f > | > \varepsilon \} < \varepsilon.$$

3

(i) Suppose that

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

is a block matrix such that A is a square matrix and $\det A \neq 0$. Show that

$$\det M = \det A \det[D - CA^{-1}B].$$

You can use the fact that for block-triangular matrices the following formulas are valid:

$$\det \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix} = \det \begin{pmatrix} X & 0 \\ Y' & Z \end{pmatrix} = \det(X)\det(Z).$$

(ii) Let x_i be the i-th column of a symmetric matrix X with the i-th entry removed, and let $X^{(i)}$ be the matrix X with column i and row i removed. Show that

$$\left[(X - zI_N)^{-1}\right]_{ii} = \frac{1}{X_{ii} - z - x_i \cdot \left(X^{(i)} - zI_{N-1}\right)^{-1} x_i},$$

where I_N and I_{N-1} denote the N-by-N and $(N-1)$-by-$(N-1)$ identity matrices, respectively.

(iii) Suppose that X_N is sequence of random symmetric N-by-N matrices, in which the diagonal entries $X_{ii} = 0$, the off-diagonal entries X_{ij} for $i < j$ are i.i.d. with $EX_{ij} = 0$ and $E\left(\sqrt{N}X_{ij}\right)^2 = 1$. Let

$$g_N(z) := \frac{1}{N}\mathrm{Tr}\,(X - zI_N)^{-1}.$$

and

$$g_N^{(i)}(z) := \frac{1}{N}\mathrm{Tr}\left(X^{(i)} - zI_{N-1}\right)^{-1}.$$

Let

$$\varepsilon_N(z) := g_N(z) - \frac{1}{-z - g_N(z)}.$$

Show that

$$\varepsilon_N(z) = \frac{1}{N}\sum_{i=1}^{N} \frac{x_i \cdot \left(X^{(i)} - zI_{N-1}\right)^{-1} x_i - g_N(z)}{(z + g_N(z))\left(z + x_i \cdot \left(X^{(i)} - zI_{N-1}\right)^{-1} x_i\right)}.$$

(iv) Show that

$$|\varepsilon_N(z)| \leqslant \frac{1}{(\mathrm{Im}z)^2}\frac{1}{N}\sum_{i=1}^{N}\left| x_i \cdot \left(X^{(i)} - zI_{N-1}\right)^{-1} x_i - g_N^{(i)}(z)\right| + O\left(\frac{1}{(\mathrm{Im}z)^3 N}\right).$$

[You can use without proof the fact that

$$|g_N^{(i)}(z) - g_N(z)| \leqslant \frac{c}{(\mathrm{Im}z)N}.]$$

END OF PAPER