Universality for non-Hermitian random matrices

Van H. Vu

Department of Mathematics
Yale University

Joint work with Terence Tao (UCLA)

M_n : random $n \times n$ matrix with (complex) iid entries

$$\textit{eigenvalues}: \ \lambda_1(M_n), \ldots, \lambda_n(M_n) \in \mathbf{C}.$$

M_n : random $n \times n$ matrix with (complex) iid entries

$$\textit{eigenvalues}: \ \lambda_1(M_n), \dots, \lambda_n(M_n) \in \mathbf{C}.$$

The *k-point correlation function*

$$\rho_n^{(k)} = \rho_n^{(k)}[M_n] : \mathbf{C}^k \to \mathbf{R}^+$$

satisfies

$$\begin{aligned} &\int_{\mathbf{C}^k} F(z_1, \dots, z_k) \rho_n^{(k)}(z_1, \dots, z_k) \, dz_1 \dots dz_k \\ &= \mathbf{E} \sum_{1 \le i_1, \dots, i_k \le n, \text{ distinct}} F(\lambda_{i_1}(M_n), \dots, \lambda_{i_k}(M_n)) \end{aligned} \tag{1}$$

M_n : random $n \times n$ matrix with (complex) iid entries

$$\textit{eigenvalues}: \; \lambda_1(M_n), \dots, \lambda_n(M_n) \in \mathbf{C}.$$

The *k-point correlation function*

$$\rho_n^{(k)} = \rho_n^{(k)}[M_n] : \mathbf{C}^k \to \mathbf{R}^+$$

satisfies

$$\begin{aligned} &\int_{\mathbf{C}^k} F(z_1, \dots, z_k) \rho_n^{(k)}(z_1, \dots, z_k) \; dz_1 \dots dz_k \\ &= \mathbf{E} \sum_{1 \le i_1, \dots, i_k \le n, \text{ distinct}} F(\lambda_{i_1}(M_n), \dots, \lambda_{i_k}(M_n)) \end{aligned} \tag{1}$$

for all continuous, compactly supported test functions F, where dz denotes Lebesgue measure on the complex plane $\mathbf{C}$.

The problem

Intuitively, $\rho_n^{(k)}(z_1, \ldots, z_k)$ (for distinct $z_1, \ldots, z_k$) is the quantity such that the probability that there is an eigenvalue of M_n in each of the disks

$$\{z : |z - z_i| \leq \epsilon\}$$

for $i = 1, \ldots, k$ is asymptotically

$$\rho_n^{(k)}(z_1, \ldots, z_k) V.$$

Normalizing factor $V =$ product of the areas of the disks.

The problem

Intuitively, $\rho_n^{(k)}(z_1, \dots, z_k)$ (for distinct $z_1, \dots, z_k$) is the quantity such that the probability that there is an eigenvalue of M_n in each of the disks

$$\{z : |z - z_i| \leq \epsilon\}$$

for $i = 1, \dots, k$ is asymptotically

$$\rho_n^{(k)}(z_1, \dots, z_k) V.$$

Normalizing factor $V =$ product of the areas of the disks.

Key Problem. Understand $\rho_n^{(k)}(z_1, \dots, z_k)$ at local scale. (How nearby eigenvalues interact ?)

Gaussian models and Universality

Gaussian model: Entries are iid standard Gaussian (complex or real).

In these cases, lots of eigenvalues statistics (global distribution, fluctuation of individual eigenvalues, correlation functions etc) can be computed explicitly.

Gaussian models and Universality

Gaussian model: Entries are iid standard Gaussian (complex or real).

In these cases, lots of eigenvalues statistics (global distribution, fluctuation of individual eigenvalues, correlation functions etc) can be computed explicitly.

Universality Phenomenon. All these statistics hold for all other "reasonable" models.

Gaussian models and Universality

Gaussian model: Entries are iid standard Gaussian (complex or real).

In these cases, lots of eigenvalues statistics (global distribution, fluctuation of individual eigenvalues, correlation functions etc) can be computed explicitly.

Universality Phenomenon. All these statistics hold for all other "reasonable" models.

We focus on the correlation functions.

Hermitian matrices.

Universality at global scale: semi-circle law (Wigner, Anorld, Pastur etc). (Tao's and Guionnet's talk)

Hermitian matrices.

Universality at global scale: semi-circle law (Wigner, Anorld, Pastur etc). (Tao's and Guionnet's talk)

Universality at local scale with respect to vague convergence:

$$\lim \int F(z_1, \ldots, z_k)\rho(z_1, \ldots, z_k)dz_1 \ldots dz_k,$$

is universal with respect to all *nice* test function F.

Hermitian matrices.

Universality at global scale: semi-circle law (Wigner, Anorld, Pastur etc). (Tao's and Guionnet's talk)

Universality at local scale with respect to vague convergence:

$$\lim \int F(z_1, \ldots, z_k)\rho(z_1, \ldots, z_k)dz_1 \ldots dz_k,$$

is universal with respect to all *nice* test function F.

Basically settled in the complex case: Universality under 2-moment assumption.

Hermitian matrices.

Universality at global scale: semi-circle law (Wigner, Anorld, Pastur etc). (Tao's and Guionnet's talk)

Universality at local scale with respect to vague convergence:

$$\lim \int F(z_1, \dots, z_k)\rho(z_1, \dots, z_k)dz_1 \dots dz_k,$$

is universal with respect to all *nice* test function F.

Basically settled in the complex case: Universality under 2-moment assumption.

Partially settled in the real case: Universality under 4-moment assumption. 2-moment if use a weaker notation of convergence (Erdos' and Tao's talks).

Non-Hermitian matrices

Non-Hermitian case. All entries are iid (mean 0 and variance 1).

Non-Hermitian matrices

Non-Hermitian case. All entries are iid (mean 0 and variance 1).

Methods used for Hermitian case usually fail.

Non-Hermitian matrices

Non-Hermitian case. All entries are iid (mean 0 and variance 1).

Methods used for Hermitian case usually fail.

Universality at global scale: circular law. (Chafai's talk).

Non-Hermitian matrices

Non-Hermitian case. All entries are iid (mean 0 and variance 1).

Methods used for Hermitian case usually fail.

Universality at global scale: circular law. (Chafai's talk).

Universality at local scale with respect to vague convergence: Nothing was known.

Complex gaussian case

Explicit calculation:

$$\rho_n^{(k)}(z_1, \ldots, z_k) = \det(K_n(z_i, z_j))_{1 \le i,j \le k} \tag{2}$$

where

$$K_n(z, w) := \frac{1}{\pi} e^{-(|z|^2+|w|^2)/2} \sum_{j=0}^{n-1} \frac{(z\overline{w})^j}{j!}. \tag{3}$$

Complex gaussian case

Explicit calculation:

$$\rho_n^{(k)}(z_1, \ldots, z_k) = \det(K_n(z_i, z_j))_{1 \le i,j \le k} \tag{2}$$

where

$$K_n(z, w) := \frac{1}{\pi} e^{-(|z|^2+|w|^2)/2} \sum_{j=0}^{n-1} \frac{(z\overline{w})^j}{j!}. \tag{3}$$

Example.

$$\rho_n^{(1)}(z) = K_n(z, z) = \frac{1}{\pi} e^{-|z|^2} \sum_{j=0}^{n-1} \frac{|z|^{2j}}{j!}. \tag{4}$$

Complex gaussian case

Explicit calculation:

$$\rho_n^{(k)}(z_1, \dots, z_k) = \det(K_n(z_i, z_j))_{1 \le i,j \le k} \tag{2}$$

where

$$K_n(z, w) := \frac{1}{\pi} e^{-(|z|^2+|w|^2)/2} \sum_{j=0}^{n-1} \frac{(z\overline{w})^j}{j!}. \tag{3}$$

Example.

$$\rho_n^{(1)}(z) = K_n(z, z) = \frac{1}{\pi} e^{-|z|^2} \sum_{j=0}^{n-1} \frac{|z|^{2j}}{j!}. \tag{4}$$

By Taylor expansion of $e^{-|z|^2}$,

$$\rho_n^{(1)}(\sqrt{n}z) \to \frac{1}{\pi} 1_{|z| \le 1}$$

for almost every $z \in \mathbf{C}$. This gives the *circular law* for complex gaussian matrices (the eigenvalues are roughly uniformly distributed in $B(0, \sqrt{n})$).

Moment matching

Definition (Independent-entry matrices)

For $k \geq 0$, we say two ensembles $M_n = (\xi_{ij})_{1 \leq i,j \leq n}$ and $M_n' = (\xi_{ij}')_{1 \leq i,j \leq n}$ have *matching moments to order k* if one has

$$\mathbf{E}\Re(\xi_{ij})^a \Im(\xi_{ij})^b = \mathbf{E}\Re(\xi_{ij}')^a \Im(\xi_{ij}')^b \tag{5}$$

whenever $1 \leq i,j \leq n$, $a, b \geq 0$ and $a + b \leq k$.

Complex case: Four Moment Theorem

Theorem

Let $M_n, \tilde{M}_n$ be (complex) random matrices matching moments with each other to fourth order. Let $z_0 \in \mathbf{C}$ be bounded and $F : \mathbf{C}^k \to \mathbf{C}$ be a nice test function. Then

$$\int_{\mathbf{C}^k} F(w_1, \ldots, w_k)\rho_n^{(k)}(\sqrt{n}z_0 + w_1, \ldots, \sqrt{n}z_0 + w_k)\, dw_1 \ldots dw_k =$$
$$\int_{\mathbf{C}^k} F(w_1, \ldots, w_k)\tilde{\rho}_n^{(k)}(\sqrt{n}z_0 + w_1, \ldots, \sqrt{n}z_0 + w_k)\, dw_1 \ldots dw_k + O(n^{-c})$$

for some absolute constant $c > 0$ (independent of k).

Complex case: Four Moment Theorem

Theorem

Let $M_n, \tilde{M}_n$ be (complex) random matrices matching moments with each other to fourth order. Let $z_0 \in \mathbf{C}$ be bounded and $F : \mathbf{C}^k \to \mathbf{C}$ be a nice test function. Then

$$\int_{\mathbf{C}^k} F(w_1, \ldots, w_k)\rho_n^{(k)}(\sqrt{n}z_0 + w_1, \ldots, \sqrt{n}z_0 + w_k)\, dw_1 \ldots dw_k =$$
$$\int_{\mathbf{C}^k} F(w_1, \ldots, w_k)\tilde{\rho}_n^{(k)}(\sqrt{n}z_0 + w_1, \ldots, \sqrt{n}z_0 + w_k)\, dw_1 \ldots dw_k + O(n^{-c}$$

for some absolute constant $c > 0$ (independent of k).

$\sqrt{n}z_0$ is a **fixed base point**

Complex case: Four Moment Theorem

Theorem

Let $M_n, \tilde{M}_n$ be (complex) random matrices matching moments with each other to fourth order. Let $z_0 \in \mathbf{C}$ be bounded and $F : \mathbf{C}^k \to \mathbf{C}$ be a nice test function. Then

$$\int_{\mathbf{C}^k} F(w_1, \ldots, w_k)\rho_n^{(k)}(\sqrt{n}z_0 + w_1, \ldots, \sqrt{n}z_0 + w_k)\ dw_1 \ldots dw_k =$$
$$\int_{\mathbf{C}^k} F(w_1, \ldots, w_k)\tilde{\rho}_n^{(k)}(\sqrt{n}z_0 + w_1, \ldots, \sqrt{n}z_0 + w_k)\ dw_1 \ldots dw_k + O(n^{-c}$$

for some absolute constant $c > 0$ (independent of k).

$\sqrt{n}z_0$ is a **fixed base point**

$$\rho_{z_0}(w_1, \ldots, w_k) := \rho_n^{(k)}(\sqrt{n}z_0 + w_1, \ldots, \sqrt{n}z_0 + w_k).$$

Real gaussian case

There are real eigenvalues. The spectrum splits:

$$\lambda_{1,\mathbf{R}}(M_n), \ldots, \lambda_{N_{\mathbf{R}}[M_n],\mathbf{R}}(M_n)$$

$$\lambda_{1,\mathbf{C}_+}(M_n), \ldots, \lambda_{N_{\mathbf{C}_+}[M_n],\mathbf{C}_+}(M_n)$$

in the upper half-plane $\mathbf{C}_+ := \{z \in \mathbf{C} : \Im(z) > 0\}$

$$\overline{\lambda_{1,\mathbf{C}_+}(M_n)}, \ldots, \overline{\lambda_{N_{\mathbf{C}_+}}(M_n)}.$$

$$N_{\mathbf{R}} + 2N_{\mathbf{C}_+} = n.$$

Mixed correlation functions

We need to work with the *mixed* correlation functions $\rho_n^{(k,l)} : \mathbf{R}^k \times \mathbf{C}_+^l \to \mathbf{R}^+$, defined for $k, l \geq 0$ by the formula

$$\int_{\mathbf{R}^k} \int_{\mathbf{C}_+^l} F(x_1, \ldots, x_k, z_1, \ldots, z_l) \rho_n^{(k,l)}(x_1, \ldots, x_k, z_1, \ldots, z_l)\ dx_i\ dz_j$$
$$= \mathbf{E} \sum_{1 \leq i_1 < \cdots < i_k \leq N_{\mathbf{R}}[M_n]} \sum_{1 \leq j_1 < \cdots < j_l \leq N_{\mathbf{C}_+}[M_n]}$$
$$F(\lambda_{i_1,\mathbf{R}}(M_n), \ldots, \lambda_{i_k,\mathbf{R}}(M_n), \lambda_{j_1,\mathbf{C}_+}(M_n), \ldots, \lambda_{j_l,\mathbf{C}_+}(M_n)). \tag{6}$$

Intuitively, one can interpret $\rho_n^{(k,l)}(x_1, \ldots, x_k, z_1, \ldots, z_l)$ (for distinct $x_1, \ldots, x_k \in \mathbf{R}$ and $z_1, \ldots, z_l \in \mathbf{C}_+$) as the unique real number such that the probability of simultaneously having an eigenvalue of M_n in each of the intervals $(x_i - \epsilon, x_i + \epsilon)$ for $i = 1, \ldots, k$ and in each of the disks $B(z_j, \epsilon)$ for $j = 1, \ldots, l$ is asymptotically

$$\rho_n^{(k,l)}(x_1, \ldots, x_k, z_1, \ldots, z_l)(2\epsilon)^k (\pi\epsilon^2)^l.$$

Intuitively, one can interpret $\rho_n^{(k,l)}(x_1, \ldots, x_k, z_1, \ldots, z_l)$ (for distinct $x_1, \ldots, x_k \in \mathbf{R}$ and $z_1, \ldots, z_l \in \mathbf{C}_+$) as the unique real number such that the probability of simultaneously having an eigenvalue of M_n in each of the intervals $(x_i - \epsilon, x_i + \epsilon)$ for $i = 1, \ldots, k$ and in each of the disks $B(z_j, \epsilon)$ for $j = 1, \ldots, l$ is asymptotically

$$\rho_n^{(k,l)}(x_1, \ldots, x_k, z_1, \ldots, z_l)(2\epsilon)^k(\pi\epsilon^2)^l.$$

In particular, $\rho^{k,0}$ reveals the correlation between the real eigenvalues.

Intuitively, one can interpret $\rho_n^{(k,l)}(x_1, \ldots, x_k, z_1, \ldots, z_l)$ (for distinct $x_1, \ldots, x_k \in \mathbf{R}$ and $z_1, \ldots, z_l \in \mathbf{C}_+$) as the unique real number such that the probability of simultaneously having an eigenvalue of M_n in each of the intervals $(x_i - \epsilon, x_i + \epsilon)$ for $i = 1, \ldots, k$ and in each of the disks $B(z_j, \epsilon)$ for $j = 1, \ldots, l$ is asymptotically

$$\rho_n^{(k,l)}(x_1, \ldots, x_k, z_1, \ldots, z_l)(2\epsilon)^k(\pi\epsilon^2)^l.$$

In particular, $\rho^{k,0}$ reveals the correlation between the real eigenvalues.

One can extend $\rho_n^{(k,l)}$ from $\mathbf{R}^k \times \mathbf{C}^l$ in a natural way.

The gaussian ensemble

The correlation functions $\rho_n^{(k,l)}$ were computed by a variety of methods (Akemann, Kanzieper,Lehmann, Sommers, Sinclair, Wieczorek, Borodin, Forerster, May etc). The case for n odd and even are *different*; for simplicity, we consider n even.

$$\rho_n^{(k,l)}(x_1, \ldots, x_k, z_1, \ldots, z_l) = \mathrm{Pf} \begin{pmatrix} \tilde{K}_n(x_i, x_{i'}) & \tilde{K}_n(x_i, z_{j'}) \\ \tilde{K}_n(z_j, x_{i'}) & \tilde{K}_n(z_j, z_{j'}) \end{pmatrix}_{1 \leq i, i' \leq k; 1 \leq j, j' \leq l}$$

$\tilde{K}_n : (\mathbf{R} \cup \mathbf{C}_+) \times (\mathbf{R} \cup \mathbf{C}_+) \to M_2(\mathbf{C})$ is a certain explicit 2×2 matrix kernel obeying the anti-symmetry law

$$\tilde{K}(\zeta, \zeta') = -\tilde{K}(\zeta', \zeta)^T. \tag{7}$$

The gaussian ensemble

Theorem (Real eigenvalues of a real gaussian matrix)

Let M_n be drawn from the real gaussian ensemble. Then

$$(Edelman - Kostlan - Shub\ 94)\ \mathbf{E} N_{\mathbf{R}}(M_n) = \sqrt{\frac{2n}{\pi}} + O(1)$$

and

$$(Forrester - Nagao\ 07)\ (Var\ N_{\mathbf{R}}(M_n) = (2-\sqrt{2})\sqrt{\frac{2n}{\pi}} + o(\sqrt{n})$$

Numerical experiment

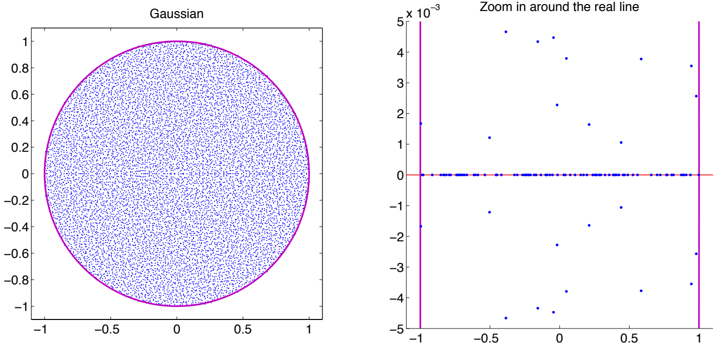

Figure: The spectrum of a random real gaussian $10,000 \times 10,000$ matrix, with additional detail near the origin to show the concentration on the real axis. Thanks to Ke Wang for the data and figure.

Real case: Four Moment Theorem

Theorem

Assume M_n and $\tilde{M}_n$ both match moments with the real gaussian matrix ensemble to fourth order. Let $k, l \geq 0$ be fixed integers, and let let x_0 and $z_0 \in \mathbf{C}$ be bounded; F, G are nice test functions. Then

$$\int_{\mathbf{R}^k}\int_{\mathbf{C}^l} F(y_1,\ldots,y_k,w_1,\ldots,w_l)\rho_n^{(k,l)}(\sqrt{n}x_0+y_1,\ldots,\sqrt{n}x_0+y_k, \sqrt{n}z_0+w_1,\ldots,\sqrt{n}z_0+w_l)\ dw_1\ldots dw_l dy_1\ldots dy_k$$

$$= \int_{\mathbf{R}^k}\int_{\mathbf{C}^l} F(y_1,\ldots,y_k,w_1,\ldots,w_l)\tilde{\rho}_n^{(k,l)}(\sqrt{n}x_0+y_1,\ldots,\sqrt{n}x_0+y_k, \sqrt{n}z_0+w_1,\ldots,\sqrt{n}z_0+w_l)\ dw_1\ldots dw_l dy_1\ldots dy_k + O(n^{-c}).$$

Real case: Four Moment Theorem

Theorem

Assume M_n and $\tilde{M}_n$ both match moments with the real gaussian matrix ensemble to fourth order. Let $k, l \geq 0$ be fixed integers, and let let x_0 and $z_0 \in \mathbf{C}$ be bounded; F, G are nice test functions. Then

$$\begin{aligned}
&\int_{\mathbf{R}^k} \int_{\mathbf{C}^l} F(y_1, \dots, y_k, w_1, \dots, w_l) \rho_n^{(k,l)}(\sqrt{n}x_0 + y_1, \dots, \sqrt{n}x_0 + y_k, \\
&\quad \sqrt{n}z_0 + w_1, \dots, \sqrt{n}z_0 + w_l)\ dw_1 \dots dw_l dy_1 \dots dy_k \\
&= \int_{\mathbf{R}^k} \int_{\mathbf{C}^l} F(y_1, \dots, y_k, w_1, \dots, w_l) \tilde{\rho}_n^{(k,l)}(\sqrt{n}x_0 + y_1, \dots, \sqrt{n}x_0 + y_k, \\
&\quad \sqrt{n}z_0 + w_1, \dots, \sqrt{n}z_0 + w_l)\ dw_1 \dots dw_l dy_1 \dots dy_k + O(n^{-c}).
\end{aligned}$$

In this theorem, we have both real and complex base points.

$$\rho_{x_0,z_0}(y_1, \dots, y_k, w_1, \dots, w_l).$$

Number of real eigenvalues

Theorem (Real eigenvalues of a real matrix)

Let M_n be real random matrix which matches moments with the real gaussian matrix ensemble to fourth order. Then

$$\mathbf{E} N_{\mathbf{R}}(M_n) = \sqrt{\frac{2n}{\pi}} + O(n^{1/2-c})$$

and

$$Var\, N_{\mathbf{R}}(M_n) = O(n^{1-c})$$

for some fixed $c > 0$. In particular, with probability $1 - o(1)$

$$N_{\mathbf{R}}(M_n) = (1 + o(1))\sqrt{\frac{2n}{\pi}}.$$

Application: Central limit theorem

Theorem (Central limit theorem, gaussian case)

(Rider 04) Let M_n be drawn from the complex gaussian ensemble. Let $r > 0$ be such that $1/r, r/n^{1/2} = o(1)$. Let $B(z_0, r)$ be a disk in the bulk of the spectrum and $N_{B(z_0,r)}$ be the number of eigenvalues of M_n in the disk. Then

$$\frac{N_{B(z_0,r)} - r^2}{r^{1/2}\pi^{-1/4}} \to N(0,1)_{\mathbf{R}}$$

Extension to general ensembles:

Theorem (Central limit theorem, general case)

Let M_n be a complex random matrix which matches moments with the complex gaussian matrix ensemble to fourth order. Then the same conclusion holds for $r \leq n^{o(1)}$.

Numerical experiments

The result must hold for all possible radii.

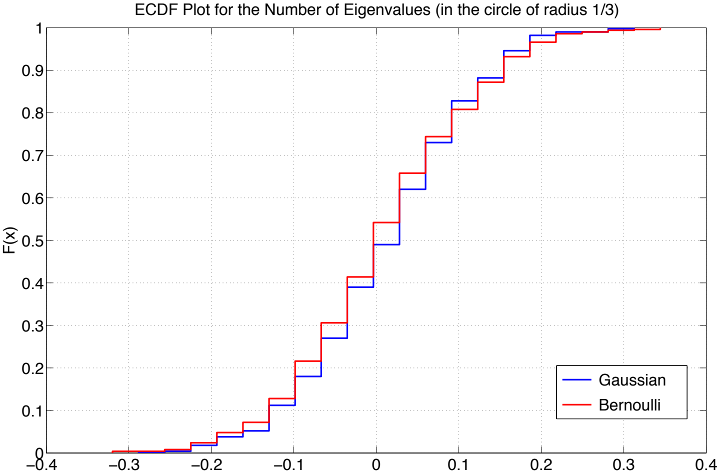

Figure: The cumulative distribution function for the number of

Application: Most eigenvalues are simple

Theorem (Most eigenvalues simple)

Let M_n be a random matrix (complex or real) which matches moments with the real or complex gaussian matrix to fourth order. Then with probability $1 - O(n^{-c})$, at most $O(n^{1-c})$ of the complex eigenvalues, and $O(n^{1/2-c})$ of the real eigenvalues, are repeated, for some fixed $c > 0$.

Conjecture. *All* eigenvalues are simple.

Difficulties

Heat flow method does not seem to work.

The eigenvalues (being mostly complex) are not ordered.

The spectrum of non-Hermitian matrices can be very sensitive to a small change.

The starting point

Instead of the Stieltjes' transform, we look at the *log-determinant*

$$\log|\det(M_n - z)|$$

for a complex number parameter z

The starting point

Instead of the Stieltjes' transform, we look at the *log-determinant*

$$\log|\det(M_n - z)|$$

for a complex number parameter z (Girko, Brown; Guionnet, Chafai's talk).

The starting point

Instead of the Stieltjes' transform, we look at the *log-determinant*

$$\log|\det(M_n - z)|$$

for a complex number parameter z (Girko, Brown; Guionnet, Chafai's talk).

Connection:

$$\log|\det(M_n - z)| = \sum_{i=1}^{n} \log|\lambda_i(M_n) - z|. \tag{8}$$

Zooming in: Jensen's formula

f analytic on $D := B(0, r)$; $a_1, a_2, ..., a_n$ are the zeros of f in the interior of D (counting multiplicity), and $f(0)0$:

$$\log|f(0)| = \sum_{i=1}^{k} \log\frac{|a_i|}{r} + \frac{1}{2\pi}\int_0^{2\pi} \log|f(re^{i\theta})|d\theta.$$

Zooming in: Jensen's formula

f analytic on $D := B(0,r)$; $a_1, a_2, ..., a_n$ are the zeros of f in the interior of D (counting multiplicity), and $f(0)0$:

$$\log|f(0)| = \sum_{i=1}^{k} \log\frac{|a_i|}{r} + \frac{1}{2\pi}\int_0^{2\pi} \log|f(re^{i\theta})|d\theta.$$

Applied Jensen's formula:

$$\log|\det(M_n - z_0)| = -\sum_{1\le i\le n:\lambda_i(M_n)\in B(z_0,r)} \log\frac{r}{|\lambda_i(M_n) - z_0|}$$

$$+\frac{1}{2\pi}\int_0^{2\pi} \log|\det(M_n - z_0 - re^{i\theta})|\ d\theta.$$

for any disk $B(z_0, r)$.

Girko's Hermitization trick

Thus (in principle, at least) information on the (joint) distribution of the log-determinants $\log|\det(M_n - z)|$ for various values of z should lead to information on the eigenvalues of M_n, and in particular on the k-point correlation functions $\rho_n^{(k)}$ of M_n.

Girko's Hermitization trick

Thus (in principle, at least) information on the (joint) distribution of the log-determinants $\log|\det(M_n - z)|$ for various values of z should lead to information on the eigenvalues of M_n, and in particular on the k-point correlation functions $\rho_n^{(k)}$ of M_n.

Girko Hermitization trick:

$$W_{n,z} := \frac{1}{\sqrt{n}} \begin{pmatrix} 0 & M_n - z \\ (M_n - z)^* & 0 \end{pmatrix} \tag{9}$$

$$\log|\det(M_n - z)| = \frac{1}{2}\log|\det W_{n,z}| + \frac{1}{2} n \log n. \tag{10}$$

Stieltjes' transform of W:

$$s_{W_{n,z}}(E+\sqrt{-1}\eta) := \frac{1}{2n}\left((W_{n,z}-E-\sqrt{-1}\eta)^{-1}\right).$$

$$\log|\det W_{n,z}| = \log|\det(W_{n,z}-\sqrt{-1}\,T)| - 2n\Im\int_0^T s_{W_{n,z}}(\sqrt{-1}\eta)\,d\eta. \tag{11}$$

Thus, in principle at least, information on the distribution of the Stieltjes transform $s_{W_{n,z}}$ will imply information on the log-determinant of $W_{n,z}$. This is the route taken, for instance, to establish the circular law (Chaifai's talk).

Trouble at the pole

Possible divergence or instability of the integral in (11) near $\eta = 0$.

Idea. Truncation; provided that one has adequate bounds on the least singular value of $W_{n,z}$.

This can be done using (Inverse-) Littlewood-Offord theorems Additive Combinatorics. One only needs a lower bound of the form n^{-100}, but much better (near optimal) bounds are known (Rudelson's talk).

Least singular value

Negative second moment identity (Tao-V. 05)

$$d_1^{-2} + \cdots + d_n^{-2} = \sigma_1^{-2} + \cdots + \sigma_n^{-2}.$$

d_i is the distance of row X_i to the hyperplane spanned by other rows.

$$d_i = |a_1\xi_1 + \ldots a_n\xi_n| = |S|,$$

where $(a_1, \ldots, a_n)$ is the (unit) normal vector and $X_i = \xi_1, \ldots, _n)$.

$$\mathbf{P}(d_i < \epsilon = \mathbf{P}(S \in (-\epsilon, \epsilon)).$$

Union bound and Monte Carlo method

Significant technical issue : formulae involving integrals require one to control the value of various random functions (log-determinants or Stieltjes transforms) for an *uncountable* number of choices of parameters such as z and η.

Can no *longer directly use union bound* to control exceptional events when the expected control on these quantities fails.

Idea. We use Monte Carlo method (Combinatorics and Theoretical compute science). This method enables us to use random sampling arguments to replace many of these integral expressions by discrete, random, approximations.

Lemma (Monte Carlo sampling lemma)

Let (X, μ) be a probability space, and let $F : X \to \mathbf{C}$ be a square-integrable function. Let $m \geq 1$, let $x_1, \ldots, x_m$ be drawn independently at random from X with distribution μ, and let S be the empirical average

$$S := \frac{1}{m}(F(x_1) + \cdots + F(x_m)).$$

Then S has mean $\int_X F \, d\mu$ and variance $\int_X (F - \int_X F \, d\mu)^2 \, d\mu$. In particular (by Chebyshev's inequality)

$$\mathbf{P}(|S - \int_X F \, d\mu| \geq \lambda) \leq \frac{1}{m\lambda^2} \int_X (F - \int_X F \, d\mu)^2 \, d\mu.$$

Monte Carlo method with large variance

If the variance is too large, the approximation is useless.

Monte Carlo method with large variance

If the variance is too large, the approximation is useless.

Smoothing. Find a *smoothing* function F_0 such that $\int_X F_0 = 0$ and $Var\,(F - F_0)$ small.

Monte Carlo method with large variance

If the variance is too large, the approximation is useless.

Smoothing. Find a *smoothing* function F_0 such that $\int_X F_0 = 0$ and $Var\ (F - F_0)$ small.

Sample $F - F_0$.

$$S := \frac{1}{m}((F - F_0)(x_1) + \cdots + (F - F_0)(x_m)).$$

The mean remains the same and variance gets smaller.

Monte Carlo method with large variance

If the variance is too large, the approximation is useless.

Smoothing. Find a *smoothing* function F_0 such that $\int_X F_0 = 0$ and $Var\ (F - F_0)$ small.

Sample $F - F_0$.

$$S := \frac{1}{m}((F - F_0)(x_1) + \cdots + (F - F_0)(x_m)).$$

The mean remains the same and variance gets smaller.

In applications, F involves the Laplacian Δ of a function with compact support; choose $F_0 = L\Delta$ where L is linear.

Four moment theorem for determinants

Theorem

Let $c_0 > 0$ be a sufficiently small absolute constant. Let M_n, M_n' be random matrices which match each other to fourth order. Let $1 \leq k \leq n^{c_0}$, let $C > 0$ be fixed, and let $z_1, \ldots, z_m \in B(0, C\sqrt{n})$. Let $G : \mathbf{R}^m \to \mathbf{C}$ be a smooth function obeying the derivative bounds

$$|\nabla^j G(x_1, \ldots, x_m)| \ll n^{c_0}$$

for all $j = 0, \ldots, 5$ and $x_1, \ldots, x_m \in \mathbf{C}$, where ∇ denotes the gradient in $\mathbf{R}^m$. Then we have

$$\begin{aligned}&\mathbf{E} G(\log|\det(M_n - z_1)|, \ldots, \log|\det(M_n - z_m)|) \\ &\quad = \mathbf{E} G(\log|\det(M_n' - z_1)|, \ldots, \log|\det(M_n' - z_m)|) + O(n^{-c_0}).\end{aligned}$$

Four moment theorem for determinants

Theorem

Let $c_0 > 0$ be a sufficiently small absolute constant. Let M_n, M_n' be random matrices which match each other to fourth order. Let $1 \le k \le n^{c_0}$, let $C > 0$ be fixed, and let $z_1, \dots, z_m \in B(0, C\sqrt{n})$. Let $G : \mathbf{R}^m \to \mathbf{C}$ be a smooth function obeying the derivative bounds

$$|\nabla^j G(x_1, \dots, x_m)| \ll n^{c_0}$$

for all $j = 0, \dots, 5$ and $x_1, \dots, x_m \in \mathbf{C}$, where ∇ denotes the gradient in $\mathbf{R}^m$. Then we have

$$\begin{aligned} &\mathbf{E} G(\log|\det(M_n - z_1)|, \dots, \log|\det(M_n - z_m)|) \\ &\quad = \mathbf{E} G(\log|\det(M_n' - z_1)|, \dots, \log|\det(M_n' - z_m)|) + O(n^{-c_0}). \end{aligned}$$

Proof. Swapping argument. The only place where four moments matter.

Error control via non-clustering

The spectrum is $B(0, \sqrt{n})$. A disk of radius $r \geq 1$ is expected to have $O(r^2)$ eigenvalues.

Error control via non-clustering

The spectrum is $B(0, \sqrt{n})$. A disk of radius $r \geq 1$ is expected to have $O(r^2)$ eigenvalues.

Non-clustering. Any disk of radius $r \geq 1$ in $B(0, 3\sqrt{n})$ has $O(r^2 n^{o(1)})$ eigenvalues, with overwhelming probability.

Error control via non-clustering

The spectrum is $B(0, \sqrt{n})$. A disk of radius $r \geq 1$ is expected to have $O(r^2)$ eigenvalues.

Non-clustering. Any disk of radius $r \geq 1$ in $B(0, 3\sqrt{n})$ has $O(r^2 n^{o(1)})$ eigenvalues, with overwhelming probability.

More is true:

Theorem (Local circular law)

Let $M_n = (\xi_{ij})_{1 \leq i,j \leq n}$ be which matches either the real or complex gaussian matrix to third order. Then for any fixed $C > 0$, one has with overwhelming probability that

$$N_{B(z_0,r)} = \int_{B(z_0,r)} \frac{1}{\pi} 1_{|z| \leq \sqrt{n}} \, dz + O(n^{o(1)} r) \qquad (12)$$

uniformly for all $z_0 \in B(0, C\sqrt{n})$ and all $r \geq 1$.

Remarks on the local law

The bound (12) is probably not best possible, even if one ignores the $n^{o(1)}$ term. In the complex gaussian case, Rider showed that the variance of $N_{B(z_0,r)}$ is $O(r)$, suggesting a fluctuation of $O(n^{o(1)} r^{1/2})$ rather than $O(n^{o(1)} r)$.

Remarks on the local law

The bound (12) is probably not best possible, even if one ignores the $n^{o(1)}$ term. In the complex gaussian case, Rider showed that the variance of $N_{B(z_0,r)}$ is $O(r)$, suggesting a fluctuation of $O(n^{o(1)}r^{1/2})$ rather than $O(n^{o(1)}r)$.

The local law was proved (about the same time) by Yau et. al. They focused on disk in the bulk of the spectrum and required only 2 matching moment. Shortly after, they extended the result to the edge, but required more moment. Very recent result of Yin (Dec 2012) claimed that one can have the edge result under 2 moment assumption as well. The methods are totally different.

Using Jensen's formula and the Monte Carlo argument, we deduce the Local Law from

Local law via Sharp concentration

Using Jensen's formula and the Monte Carlo argument, we deduce the Local Law from

Theorem (Concentration bound on log-determinant)

Let M_n be a random matrix matching the real or complex gaussian ensemble to third order. Then for any fixed $C > 0$, and any $z_0 \in B(0, C)$, with overwhelming probability

$$\log|\det(M_n - z_0\sqrt{n})| = \frac{1}{2} n \log n + \frac{1}{2} n(|z_0|^2 - 1) + n^{o(1)} \ \textit{for}\ |z_0| \leq 1$$

$$\log|\det(M_n - z_0\sqrt{n})| = \frac{1}{2} n \log n + n \log|z_0| + n^{o(1)}, \ \textit{for}\ |z_0| \geq 1.$$

Toy case. $z_0 = 0$;

$$\log|\det M_n| = \frac{1}{2}n\log n - \frac{1}{2}n + n^{o(1)}.$$

Toy case. $z_0 = 0$;

$$\log|\det M_n| = \frac{1}{2} n \log n - \frac{1}{2} n + n^{o(1)}.$$

The upper bound is easy, via Turán's identity $\mathbf{E}|\det M_n|^2 = n! \approx (n/e)^n$.

Toy case. $z_0 = 0$;

$$\log|\det M_n| = \frac{1}{2} n \log n - \frac{1}{2} n + n^{o(1)}.$$

The upper bound is easy, via Turán's identity
$\mathbf{E}|\det M_n|^2 = n! \approx (n/e)^n$.
Using Swapping method, it suffices to verify the gaussian case

Theorem

The statement holds when M_n is drawn from the real or complex gaussian ensemble.

Tridiagonalisation of GUE

Tridiagonalisation of GUE matrices (Trotter) Edelman's talk:

$$M_n' = \begin{pmatrix} a_1 & b_1 & 0 & \dots & 0 & 0 \\ b_1 & a_2 & b_2 & \dots & 0 & 0 \\ 0 & b_2 & a_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & a_{n-1} & b_{n-1} \\ 0 & 0 & 0 & \dots & b_{n-1} & a_n \end{pmatrix}$$

$a_1, \dots, a_n \equiv N(0,1)_{\mathbf{R}}$ being standard real Gaussians, and each b_i having a χ-distribution (all independent)

$$b_i = (\sum_{j=1}^{i} |z_{i,j}|^2)^{1/2}$$

where $z_{i,j} \equiv N(0,1)_{\mathbf{C}}$ are iid complex gaussians. Then the joint eigenvalue distribution of GUE is identical to the joint eigenvalue distribution of M_n'.

Hessenberg form

Lemma

Let M_n be a complex gaussian matrix, and let M'_n be the random matrix

$$M'_n = \begin{pmatrix} \xi_{11} & \chi_{n-1,\mathbf{C}} & 0 & 0 & \dots & 0 \\ \xi_{21} & \xi_{22} & \chi_{n-2,\mathbf{C}} & 0 & \dots & 0 \\ \xi_{31} & \xi_{32} & \xi_{33} & \chi_{n-3,\mathbf{C}} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \xi_{(n-1)1} & \xi_{(n-1)2} & \xi_{(n-1)3} & \xi_{(n-1)4} & \dots & \chi_{1,\mathbf{C}} \\ \xi_{n1} & \xi_{n2} & \xi_{n3} & \xi_{n4} & \dots & \xi_{nn} \end{pmatrix}$$

where ξ_{ij} for $1 \leq j \leq i \leq n$ are iid copies of the complex gaussian $N(0,1)_{\mathbf{C}}$, and for each $1 \leq i \leq n-1$, $\chi_{i,\mathbf{C}}$ is a complex χ distribution of i degrees of freedom. The the two matrices have the same spectrum distribution.

First showed by Krisnapur and Virag (2011).

Lower triangular form

The next step is to apply a proper linear operator (*preserving the determinant*) to bring $M'_n - z_0$ into a lower triangular matrix. The diagonal entries of this matrix are

$$(\sqrt{|a_1|^2 + \chi^2_{n-1,\mathbf{C}}}, \sqrt{|a_2|^2 + \chi^2_{n-2,\mathbf{C}}}, \dots, \sqrt{|a_{n-1}|^2 + \chi^2_{1,\mathbf{C}}}, a_n).$$

$$a_1 := N(0,1)_{\mathbf{C}} - z_0\sqrt{n} \tag{13}$$

$$a_{i+1} := \frac{-z_0\sqrt{n}a_i}{\sqrt{|a_i|^2 + \chi^2_{n-i,\mathbf{C}}}} + N(0,1)_{\mathbf{C}}. \tag{14}$$

Lower triangular form

The next step is to apply a proper linear operator (*preserving the determinant*) to bring $M'_n - z_0$ into a lower triangular matrix. The diagonal entries of this matrix are

$$(\sqrt{|a_1|^2 + \chi^2_{n-1,\mathbf{C}}}, \sqrt{|a_2|^2 + \chi^2_{n-2,\mathbf{C}}}, \dots, \sqrt{|a_{n-1}|^2 + \chi^2_{1,\mathbf{C}}}, a_n).$$

$$a_1 := N(0,1)_{\mathbf{C}} - z_0\sqrt{n} \tag{13}$$

$$a_{i+1} := \frac{-z_0\sqrt{n}a_i}{\sqrt{|a_i|^2 + \chi^2_{n-i,\mathbf{C}}}} + N(0,1)_{\mathbf{C}}. \tag{14}$$

Key. Show that the a_i behave as expected, with high probability.

Divide and Conquer Martingale

Let $Y = Y(\xi_1, \ldots, \xi_n)$ be a random variable depending on independent atom variables $\xi_i \in \mathbf{C}$. For $1 \leq i \leq n$ and $\xi = (\xi_1, \ldots, \xi_n) \in \mathbf{C}^n$, define the martingale differences

$$C_i(\xi) := |\mathbf{E}(Y|\xi_1, \ldots, \xi_i) - \mathbf{E}(Y|\xi_1, \ldots, \xi_{i-1})|.$$

Azuma (1960s): $C_i \leq \alpha_i$ with probability one, then

$$\mathbf{P}\left(|Y - \mathbf{E}Y| \geq \lambda \sqrt{\sum_{i=1}^{n} \alpha_i^2}\right) \leq \exp(-c\lambda^2).$$

Divide and Conquer Martingale

Let $Y = Y(\xi_1, \ldots, \xi_n)$ be a random variable depending on independent atom variables $\xi_i \in \mathbf{C}$. For $1 \leq i \leq n$ and $\xi = (\xi_1, \ldots, \xi_n) \in \mathbf{C}^n$, define the martingale differences

$$C_i(\xi) := |\mathbf{E}(Y|\xi_1, \ldots, \xi_i) - \mathbf{E}(Y|\xi_1, \ldots, \xi_{i-1})|.$$

Azuma (1960s): $C_i \leq \alpha_i$ with probability one, then

$$\mathbf{P}\left(|Y - \mathbf{E}Y| \geq \lambda \sqrt{\sum_{i=1}^{n} \alpha_i^2}\right) \leq \exp(-c\lambda^2).$$

In applications, the assumption that $C_i \leq \alpha_i$ with probability one usually fails.

Divide and Conquer Martingale

Let $Y = Y(\xi_1, \ldots, \xi_n)$ be a random variable depending on independent atom variables $\xi_i \in \mathbf{C}$. For $1 \le i \le n$ and $\xi = (\xi_1, \ldots, \xi_n) \in \mathbf{C}^n$, define the martingale differences

$$C_i(\xi) := |\mathbf{E}(Y|\xi_1, \ldots, \xi_i) - \mathbf{E}(Y|\xi_1, \ldots, \xi_{i-1})|.$$

Azuma (1960s): $C_i \le \alpha_i$ with probability one, then

$$\mathbf{P}\left(|Y - \mathbf{E}Y| \ge \lambda\sqrt{\sum_{i=1}^n \alpha_i^2}\right) \le \exp(-c\lambda^2).$$

In applications, the assumption that $C_i \le \alpha_i$ with probability one usually fails. We have this refinement:

$$\mathbf{P}\left(|Y - \mathbf{E}Y| \ge \lambda\sqrt{\sum_{i=1}^n \alpha_i^2}\right) \le \exp(-c\lambda^2) + \sum_{i=1}^n \mathbf{P}(C_i(\xi) \ge \alpha_i).$$

Open questions

Reduce the number of moments ?

Open questions

Reduce the number of moments ?

Gauss divisible models ?

Open questions

Reduce the number of moments ?

Gauss divisible models ?

Distribution of the "largest" eigenvalue ? (known in the gaussian case; Rider).

Open questions

Reduce the number of moments ?

Gauss divisible models ?

Distribution of the "largest" eigenvalue ? (known in the gaussian case; Rider).

Distribution of the permanent ?