Products of Independent Random Matrices

- Sean O'Rourke & A. S.
  Elec. J. Probab. 2011

- Sean O'Rourke,
  David Renfrew,
  A. S. & Van Vu
  arXiv: 1403.6080 [math.PR]

Three basic ensembles of random matrices

(i) Ginibre ensemble

$A = (a_{jk})_{j,k=1}^{N}$ is an $N \times N$ matrix with i.i.d. standard complex (real) Gaussian entries

(ii) GUE (GOE)

$$W = \frac{A + A^*}{\sqrt{2}} \quad \left(W = \frac{A + A^t}{\sqrt{2}} \text{ in the real case} \right)$$

(iii)

Elliptic (Real) Gaussian Matrices

$$M = \cos\theta \, A + \sin\theta \, A^t .$$

$$\{ (m_{ij}, m_{ji}), \ 1 \le i < j \le N \} \cup \{ m_{kk}, \ 1 \le k \le N \}$$

are independent (Gaussian) random elements.

$\theta = 0$ corresponds to the Ginibre ensemble

$\theta = \frac{\pi}{4}$ corresponds to the GOE.

Global Eigenvalue Distribution

(i) Ginibre

In the complex case the joint probability distribution of the eigenvalues is given by the density

$$p_N(z_1, \ldots, z_N) = \frac{1}{\pi^N \prod_{m=1}^{N} m!} \exp\left(-\sum_{k=1}^{N} |z_k|^2\right) \prod_{i<j} |z_i - z_j|^2$$

Mehta used the joint density to compute the limiting spectral measere of the complex Ginibre ensemble

Namely, the spectral density for finite N is known explicitely.

$$\rho_N(z) = \frac{1}{\pi} \exp\left(-|z|^2\right) \sum_{j=0}^{N-1} \frac{|z|^{2j}}{j!}$$

In particular,

$$\rho_N(z) \longrightarrow \begin{cases} \frac{1}{\pi} & \text{if } |z| < (1-\delta)\sqrt{N} \\ 0 & \text{if } |z| > (1+\delta)\sqrt{N} \end{cases}$$

so the eigenvalues of $\frac{1}{\sqrt{N}} A$ are distributed uniformly on the unit disk in the limit $N \to \infty$.

This result is known as the circular law in the special Gaussian (complex Ginibre) case

In this case it was proved by M. L. Mehta in 1967.

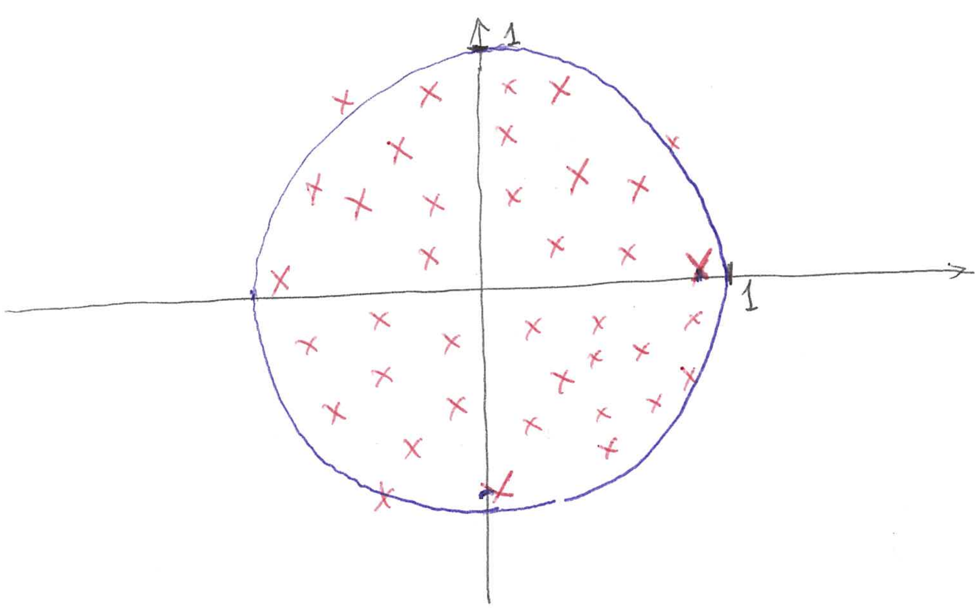
Mehta heavily relied on the formula for the joint distribution of the eigenvalues (*) discovered by Ginibre two years earlier.

It took more than 40 years to extend the circular law to the non-Gaussian case in the full generality

The result is known as the circular law.

V. Girko, Z.D. Bai, F. Götze & A. Tikhomirov, G. Pan & W. Zhou, T. Tao & V. Vu, ...

extended the result to a non-Gaussian setting. In particular, Tao & Vu proved it assuming that the second moment is bounded (2010).



the eigenvalues of $\frac{1}{\sqrt{N}} A$

Now let us briefly consider the real symmetric ensemble.

$$W_N = \frac{A + A^t}{\sqrt{2}}$$

The eigenvalues of $\frac{W_N}{\sqrt{N}}$ are real random variables

$$\lambda_1, \lambda_2, \ldots, \lambda_N$$

and their empirical distribution function $F_N(x) = \frac{1}{N} \# \{ i : \lambda_i \leq x \}$

converges in the limit $N \to \infty$

to the Wigner Semicircle Law

given by the density

$$\rho(x) = \begin{cases} \frac{1}{2\pi}\sqrt{4-x^2} & \text{if} \quad |x| \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

The proof in the non-Gaussian case is not that much more difficult than in the Gaussian case

(unlike in the complex case!)

Finally, in the elliptic ensemble $M = \cos\theta\, A + \sin\theta\, A^t$.

the eigenvalues of $\frac{1}{\sqrt{N}} M$

uniformly fill the ellipsoid

$$\mathcal{E}_\rho = \left\{ z \in \mathbb{C} : \frac{(\operatorname{Re} z)^2}{(1+\rho)^2} + \frac{(\operatorname{Im} z)^2}{(1-\rho)^2} \leq 1 \right\},$$

$$\rho = \mathbb{E}\, m_{ij} m_{ji} = \sin(2\theta), \qquad i \neq j$$

in the limit $N \to \infty$

(the non-Gaussian case proved by H. Nguyen & S. O'Rourke in 2012).

Now let us talk about products of non-Hermitian random matrices.

Let $M_N = N^{-\frac{m}{2}} \prod_{i=1}^{m} A_{N,i}$, where

$m > 1$ is fixed and

$A_{N,1}, A_{N,2}, \ldots, A_{N,m}$ are independent random matrices. For each

$1 \leq k \leq m$ the entries of $A_{N,k}$

are i.i.d. copies of a complex random variable ξ_k with mean zero and unit variance.

In addition, we require that

$\mathbb{E} |\xi_k|^{2+\eta} < \infty$ for some $\eta > 0$.

Theorem
(S. O'Rourke & A.S.; C. Bordenave; F. Götze & A. Tikhomirov 2011)

The empirical distribution function of the eigenvalues of $M_N = N^{-\frac{m}{2}} \prod_{i=1}^{N} A_{N,i}$

converges weakly almost surely to the limiting distribution supported on the unit disk

with the density

$$f_m(z) := \begin{cases} \dfrac{1}{m\pi} |z|^{\frac{2}{m}-2} & \text{for } |z| \leq 1 \\[2em] 0 & \text{for } |z| > 1. \end{cases}$$

Recently, we have been able to extend the result to products of elliptic and Wigner random matrices.

Definition

Let (ξ_1, ξ_2) be a random vector in $\mathbb{R}^2$ and let η be a real random variable.

We say that $Y_N = (y_{ij})_{i,j=1}^{N}$ is a $N \times N$ real elliptic random matrix with atom variables (ξ_1, ξ_2), η if

- (independence)

$$\{ y_{ii} : 1 \le i \le N \} \cup \{ (y_{jk}, y_{kj}) : 1 \le j < k \le N \}$$

is a collection of independent random elements.

- (off-diagonal entries)

$$\{ (y_{jk}, y_{kj}) : 1 \le j < k \le N \}$$

is a collection of i.i.d. copies of (ζ_1, ζ_2).

- (diagonal entries)

$\{ y_{ii} : 1 \le i \le N \}$ is a collection of i.i.d. copies of η.

Theorem 1 (S. O'Rourke, D. Renfrew, A.S., and V. Vu)

Let $m > 1$ be a fixed integer.

For each $1 \le k \le m$, let

$(\xi_{k,1}, \xi_{k,2})$, η_k be real random

elements satisfying

(i) $\xi_{k,1}$, $\xi_{k,2}$ both have mean zero and unit variance.

(ii) $\mathbb{E}|\xi_{k,1}|^{2+\varepsilon} + \mathbb{E}|\xi_{k,2}|^{2+\varepsilon} < \infty$

for some $\varepsilon > 0$.

(iii) $\rho_k := \mathbb{E}[\xi_{k,1}\xi_{k,2}]$ satisfies

$|\rho_k| < 1$ if $m > 2$

(DO NOT NEED IT FOR $m = 2$!)

(iv) η_k has mean zero and finite variance.

For each $N \geq 1$ and $1 \leq k \leq m$ let $Y_{N,k}$ be an $N \times N$ real elliptic random matrix with atom variables $(\xi_{k,1}, \xi_{k,2}), \eta_k$. Assume that $Y_{N,1}, \ldots, Y_{N,m}$ are independent. Then the ESD of the product

$N^{-\frac{m}{2}} \prod_{k=1}^{N} Y_{N,k}$ converges almost surely to the distribution supported on $|z| \leq 1$

with the density

$$f_m(z) = \frac{1}{m\pi} |z|^{\frac{2}{m}-2}, \qquad |z| \le 1.$$

Remark The result also holds in the case

$$M_N = N^{-\frac{m}{2}} \prod_{k=1}^{m} \left(Y_{N,k} + B_{N,k} \right)$$

provided $B_{N,k}$ is a $N \times N$ deterministic matrix s.t.

$$\max_{1 \le k \le m} \operatorname{rank}(B_{N,k}) = O(N^{1-\varepsilon})$$

and

$$\sup_{N \ge 1} \max_{1 \le k \le m} \frac{1}{N^2} \| B_{N,k} \|_2 < \infty .$$

Key Steps

I) Linearization

Instead of $N \times N$ matrix M_N one considers $mN \times mN$ block matrix of the form

$$Z_N = \frac{1}{\sqrt{N}} \begin{bmatrix} 0 & Y_{N,1} & 0 & \cdots & 0 \\ 0 & 0 & Y_{N,2} & \cdots & 0 \\ \vdots & & & & \vdots \\ & & & & 0 \\ 0 & 0 & \cdots & & Y_{N,m-1} \\ Y_{N,m} & 0 & \cdots & & 0 \end{bmatrix}$$

The result follows if one can prove that the ESD of Z_N converges to the circular law in the limit $N \to \infty$

II) Hermitization

Let M be a complex (random) matrix. The Stieltjes transform of the ESD of M is defined as

$$s_N(z) := \frac{1}{N} \sum_{j=1}^{N} \frac{1}{\lambda_j(M) - z}$$

$$= \int_{\mathbb{C}} \frac{1}{\lambda - z} \, dF^M(s,t), \quad \text{where}$$

$\lambda = s + it$, and $z = x + iy$.

Since $s_N(z)$ is analytic everywhere except at the poles (the eigenvalues of M), the real part of $s_N(z)$ uniquely determines the eigenvalues

$$\mathrm{Re}\ s_N(z) = \frac{1}{N}\sum_{j=1}^{N} \frac{\mathrm{Re}\ \lambda_j(M) - x}{|\lambda_j - z|^2}$$

$$= -\frac{1}{2N}\sum_{j=1}^{N} \frac{\partial}{\partial x} \log|\lambda_j(M) - z|^2$$

$$= -\frac{1}{2}\frac{\partial}{\partial x}\frac{1}{N}\sum_{j=1}^{N} \log|\lambda_j(M) - z|^2$$

$$= -\frac{1}{2}\frac{\partial}{\partial x}\frac{1}{N}\sum_{j=1}^{N} \log \mu_j$$

where $\mu_1, \dots, \mu_N$ are the eigenvalues of $(M - z\,\mathrm{Id})(M - z\,\mathrm{Id})^*$.

Thus

$$\mathrm{Re}\ s_N(z) = -\frac{1}{2}\frac{\partial}{\partial x}\int_0^\infty \log t \ dF^{(M-zI)(M-zI)^*}(t)$$

Asymptotic Equipartition Property and CLT for Beta Random Matrix Ensembles

(joint with A. Bufetov, S. Mkrtchyan, and M. Shcherbina)

Consider a probability distribution in $\mathbb{R}^N$ of the form

$$P_N^{\beta}(\lambda_1, \dots, \lambda_N) = \frac{1}{Z_N(\beta)} \prod_{1 \le i < j \le N} |\lambda_i - \lambda_j|^{\beta} \cdot \prod_{k=1}^{N} e^{-\beta N V(\lambda_i)/2} .$$

where $\beta > 0$ and the potential V is a real analytic function satisfying

$$V(\lambda) \geq 2(1+\varepsilon)\log(1+|\lambda|)$$

for all sufficiently large λ.

Such ensembles are known as **beta ensembles** in Random Matrix Theory.

Examples of Beta Ensembles

Joint probability density of the eigenvalues of unitary (orthogonal, symplectic) ensemble of random matrices

$$P(dM) = \text{const}_N(\beta)\, e^{-\frac{\beta}{2} N \operatorname{Tr} V(M)}\, dM$$

where dM is the Lebesgue measure on the space of $N \times N$ Hermitian $(\beta = 2)$, real symmetric $(\beta = 1)$, or self-dual quaternion Hermitian matrices $(\beta = 4)$.

It is known (Pastur & Shcherbina, Johansson, Borot & Guionnet) that if the potential V is sufficiently smooth (V' is Hölder continuous) then there exists an equilibrium measure μ^V, which is absolutely continuous with respect to the Lebesgue measure and has compact support.

The equilibrium measure maximizes the functional

$$\mathcal{E}_V(\mu) := \iint \log|x-y| \, d\mu(x) \, d\mu(y)$$

$$- \int V(x) \, d\mu(x)$$

on the space of the probability measures on $\mathbb{R}$.

The support of μ^V and the density ρ^V are uniquely determined by the Euler-Lagrange variational equations

Euler-Lagrange variational equations

$$\int \log|x-y| \, d\mu^V(y) - V(x) \le \ell, \quad x \in \mathbb{R}$$

$$\int \log|x-y| \, d\mu^V(y) - V(x) = \ell, \quad x : \rho^V(x) > 0$$

In the quadratic case $V(x) = \frac{x^2}{2}$ the equilibrium measure is the Wigner semicircle distribution

$$\mu(dx) = \frac{1}{2\pi} \sqrt{4-x^2} \, \mathbb{1}_{[-2,2]}(x) \, dx.$$

Consider the βeta ensemble

$$P_N^\beta(\lambda_1, \dots, \lambda_N) = \frac{1}{Z_N(\beta)} \prod_{i<j} |\lambda_i - \lambda_j|^\beta \cdot \prod_{k=1}^{N} e^{-\frac{\beta}{2} N V(\lambda_k)} \qquad (*)$$

and define the random variable

$$X_N(\bar{\lambda}) := -\frac{\log P_N^\beta(\bar{\lambda})}{N},$$ where

$\bar{\lambda} = (\lambda_1, \dots, \lambda_N)$, is distributed according to the probability measure $(*)$.

Our first result states the asymptotic equipartition property for the ensemble.

Theorem

Let V be a real analytic function growing faster than $\log(1+\lambda^2)$ as $|\lambda| \to \infty$ whose equilibrium density has q-interval support σ $(q \geq 1)$. Also assume that ρ^V is generic (i.e. $\rho^V \neq 0$ in the interval points of σ, behaves like square root near the edges of σ, and the function

$$v(x) := \int \log|x-y|\, d\mu^V(y) - V(x)$$

attains its maximum only if $x \in \sigma$). Then for any $\beta > 0$ the random variable $X_N(\bar{\lambda}) = -\frac{\log P_N^\beta(\bar{\lambda})}{N}$ converges almost surely to some constant $E_\beta(V)$.

Remark

Analogous results for the Plancherel measure on the set of partitions of n was proved by A. Bufetov.

Central Limit Theorem

Consider the potential energy

$$H_N(\bar{\lambda}) = N \sum_{i=1}^{N} V(\lambda_i) - \sum_{i \neq j} \log|\lambda_i - \lambda_j|.$$

Note that

$$P_N^\beta(\bar{\lambda}) = \frac{1}{Z_N(\beta)} e^{-\frac{\beta}{2} H_N(\bar{\lambda})}.$$

What can be said about fluctuations of H_N ?

Theorem

Let $$H_N(\bar{\lambda}) = N \sum_{i=1}^{N} V(\lambda_i) - \sum_{i \neq j} \log|\lambda_i - \lambda_j|$$

Then

$$\frac{H_N(\bar{\lambda}) - C_{N,\beta}}{\sqrt{N}} \xrightarrow[N\to\infty]{d} N(0, Var(\beta)),$$

where $C_{N,\beta}$ is an appropriate centering constant,

$$Var(\beta) = \frac{2}{\beta} - \psi'\left(1 + \frac{\beta}{2}\right),$$ and

ψ is the digamma function

$$\psi(x) = \frac{d}{dx} \log \Gamma(x) = \frac{\Gamma'(x)}{\Gamma(x)}.$$

Remark

It is important to note that while the fluctuation of the linear statistic

$$\sum_{i=1}^{N} V(\lambda_i)$$

is asymptotically Gaussian in the one-cut case, it is non-Gaussian, in general, in the multi-cut case (L. Pastur 2006).

Yet the fluctuation of

$$N \sum V(\lambda_i) - \sum_{i \neq j} \log |\lambda_i - \lambda_j|$$

is always Gaussian.

Remark

Both terms $N \sum_{i=1}^{N} V(\lambda_i)$.

and $\sum_{1 \le i \neq j \le N} \log |\lambda_i - \lambda_j|$

in the expression for potential energy have fluctuations of order N. At the same time, their difference fluctuates on a much smaller scale, namely $\sqrt{N}$.

The cancellations take place because of the Euler-Lagrange variational equation for μ^V.