High trace methods in random matrix theory

Charles Bordenave

CNRS & Aix-Marseille Université

Introduction

Spectral theorem

Let $H \in M_n(\mathbb{C})$ be an Hermitian matrix.

There is an orthonormal basis of eigenvectors of H. Its eigenvalues (counting multiplicities) are real and are denoted by

$$\lambda_n \leqslant \cdots \leqslant \lambda_1.$$

The spectrum is the set

$$\sigma(H) = \{\lambda_i : i = 1, \cdots, n\} = \{\lambda : H - \lambda I \text{ is not invertible}\}.$$

The operator norm coincides with the spectral radius

$$\|H\| = \sup_{f \in \mathbb{R}^n} \frac{\|Hf\|_2}{\|f\|_2} = \max\{|\lambda_i| : i = 1, \cdots, n\}.$$

Trace formula

For any analytic function f,

$$\operatorname{Tr} f(H) = \sum_{x=1}^{n} f(H)_{xx} = \sum_{i=1}^{n} f(\lambda_i).$$

The above formula identifies sum of diagonal entries $\sum_x f(H)_{xx}$ and linear statistics of eigenvalues $\sum_i f(\lambda_i)$. It connects a geometric information with a spectral information.

In many situations, for good test functions f, the entries of $f(H)$ can be computed or estimated.

From the spectral theorem, we may also retrieve information on eigenvectors from the individual entries $f(H)_{xy}$.

Resolvent versus Polynomial

For $f(\lambda) = \lambda^k$

$$(H^k)_{xy} = \sum \prod_{t=1}^{k} H_{x_{t-1}x_t}$$

where the sum is over all $(x_0, \ldots, x_k) \in \{1, \ldots, n\}^{k+1}$ such that $x_0 = x$, $x_k = y$: paths of length k from x to y.

For $f(\lambda) = (\lambda - z)^{-1}$

$$f(H) = (H - z)^{-1}$$

is the resolvent of H at $z \in \mathbb{C} \backslash \sigma(H)$.

The resolvent is (essentially) the generating function of the powers: formally

$$(H - z)^{-1} = \sum_{k=0}^{\infty} z^{-k-1} H^k.$$

Resolvent versus Polynomial

In (non-integrable) random matrix theory, the study of resolvent is made thanks to analytical and probabilistic methods. The study of polynomials relies mostly combinatorial arguments.

Resolvent methods have proven to be more powerful but not adapted to all random matrix models.

Roadmap

In this course, we will investigate random matrices thanks to their polynomials.

Two classes of random Hermitian matrices considered

- ⋆ Wigner: independent coefficients above the diagonal.
- ⋆ Uniform regular: for example, adjacency matrix of random regular graphs or matrices obtained from uniform permutation matrices.

Empirical spectral distribution (ESD)

The empirical distribution of eigenvalues / empirical spectral distribution / spectral measure / density of states is the probability measure on $\mathbb{R}$

$$\mu_H = \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i}.$$

Hence,

$$\frac{1}{n} \operatorname{Tr} f(H) = \int f(\lambda) d\mu_H(\lambda).$$

Large dimensional matrices

We will be interested in sequences of matrices $H = H_n$.

First basic question: for some probability measure μ,

$$\lim_{n\to\infty} \mu_H = \mu \quad \text{(for the weak convergence topology).}$$

If H is random, μ_H is a random probability measure. The convergence holds in probability / a.s. if for all bounded continuous functions, in probability / a.s., $\int \varphi d\mu_H$ converges.

Second basic question: for some reals $a \geqslant b$,

$$\lim_{n\to\infty} \lambda_1 = a \quad \text{and} \quad \lim_{n\to\infty} \lambda_n = b.$$

Classical (high) trace method

Method of moments

Definition

Let μ be a probability measure on $\mathbb{R}$ with all moments $m_k = \int \lambda^k d\mu$ finite, $k \in \mathbb{N}$.

It is uniquely characterized by its moments (UCM) if μ is the unique measure with moment sequence $(m_k), k \in \mathbb{N}$.

Method of moments

Provided that, for some $a > 0$ and all k (even, large enough),

$$|m_k| < (ak)^k,$$

the Fourier transform analytic in a neighborhood of 0:

$$\int e^{it} d\mu = \sum_{k=0}^{\infty} \frac{(it)^k}{k!} m_k.$$

In particular μ is UCM.

If μ has support in $[-a, a]$, then $|m_k| \leqslant a^k$ and μ is UCM. Other examples: sub-Gaussian or sub-exponential variables.

Carleman's condition (1922): μ is UCM if

$$\sum_{k=0}^{\infty} m_{2k}^{-\frac{1}{2k}} = \infty.$$

Method of moments

Lemma (Method of moments)

Let μ be UCM and μ_n be a sequence of probability measures on $\mathbb{R}$. If for all $k \in \mathbb{N}$,

$$\lim_{n\to\infty} \int \lambda^k d\mu_n = m_k.$$

Then μ_n converges to μ weakly.

Corollary (Trace method)

Let μ be UCM and $H = H_n \in M_n(\mathbb{C})$ be a sequence of Hermitian matrices. If for all $k \in \mathbb{N}$,

$$\lim_{n\to\infty} \frac{1}{n} \operatorname{Tr} H^k = m_k.$$

Then μ_H converges to μ weakly.

High trace and operator norm

Recall that $\|H\| = \max |\lambda_i|$. For any even k,

$$\frac{1}{n}\mathrm{Tr}(H^k) \leqslant \|H\|^k \leqslant \mathrm{Tr}(H^k)$$

.

If $k \gg \log n$ then $n^{1/k} \to 1$ and

$$\|H\| \sim \mathrm{Tr}(H^k)^{1/k}.$$

Lower bound on largest eigenvalue

Assume that μ_H converges weakly to μ with support $[b, a]$, $|b| \leqslant a$.

Then, for any $\varepsilon > 0$, $\mu(a - \varepsilon, \infty) > 2c_\varepsilon > 0$ and for all n large enough

$$|\{i : \lambda_i > a - \varepsilon\}| = n\mu_H(a - \varepsilon, \infty) > nc_\varepsilon.$$

Notably,

$$\liminf_n \lambda_1 \geqslant a.$$

Also, for any $\varepsilon > 0$, for all even k large enough, for all n large enough,

$$\operatorname{Tr} H^k \geqslant nc_\varepsilon(a - \varepsilon)^k.$$

Lower bound on largest eigenvalue

Assume that μ_H converges weakly to μ with support $[b, a]$, $|b| \leqslant a$.

Then, for any $\varepsilon > 0$, $\mu(a-\varepsilon, \infty) > 2c_\varepsilon > 0$ and for all n large enough

$$|\{i : \lambda_i > a - \varepsilon\}| = n\mu_H(a-\varepsilon, \infty) > nc_\varepsilon.$$

Notably,

$$\liminf_n \lambda_1 \geqslant a.$$

Also, for any $\varepsilon > 0$, for all even k large enough, for all n large enough,

$$\operatorname{Tr} H^k \geqslant n(a-\varepsilon)^k.$$

Upper bound on largest eigenvalue

Conversely, assume that for even k,

$$\sum_{i=1}^{n} \lambda_i^k = \mathrm{Tr}(H^k) \leqslant n(a+\varepsilon)^k.$$

Since $\lambda_1^k \leqslant \mathrm{Tr}(H^k)$, we find,

$$\lambda_1 \leqslant n^{\frac{1}{k}}(a+\varepsilon).$$

Assume that H is random. If for some $k = k_n \gg \log n$ and $\varepsilon = \varepsilon_n \to 0$

$$\mathbb{E}\mathrm{Tr}(H^k) \leqslant n(a+\varepsilon)^k,$$

then $\limsup_n (\mathbb{E}\lambda_1^k)^{\frac{1}{k}} \leqslant a$. In particular, a.s. $\limsup_n \lambda_1 \leqslant a$.

In Summary

For a sequence of random matrices $H = H_n$,

For all fixed k, the convergence of

$$\frac{1}{n}\mathrm{Tr}(H^k)$$

will imply the convergence of μ_H.

For large $k \gg \log n$, estimates on

$$\mathbb{E}\mathrm{Tr}(H^k)$$

turn into estimates on the operator norm.

Remarks

These bounds can be turned into quantitative bounds.

Variants are possible, for example, replace the monomials H^k by another sequence of well-chosen polynomials.

Wigner seminal work on random matrices (1955) used the trace method for fixed k.

Fűredi and Komlós (1981) introduced the high trace method, $k \gg 1$ in random matrix theory.

Semicircle law for Wigner matrices

Sparse Wigner matrices

Let $H \in M_n(\mathbb{C})$ be an Hermitian random matrix with independent centered entries $(H_{xy})_{x \geqslant y}$ above the diagonal,

$$\text{for all } x \neq y, \quad \mathbb{E}|H_{xy}|^2 = \frac{1}{n} \quad \text{and} \quad a.s.\!-\max_{x,y} |H_{xy}| \leqslant \frac{1}{q}.$$

The scalar q controls the sparsity of H.

Erdős-Rényi random graph

Let $A \in M_n(\mathbb{R})$ be a symmetric matrix with iid Bernoulli entries $(A_{xy})_{x \geqslant y}$ above the diagonal, $A_{xx} = 0$.

$$\mathbb{P}(A_{xy} = 1) = 1 - \mathbb{P}(A_{xy} = 0) = \frac{d}{n}.$$

Then d is (asymptotically) the average degree of a vertex

$$d - \frac{d}{n} = \mathbb{E} \sum_{y=1}^{n} A_{xy} = \mathbb{E} \deg(x) = \mathbb{E} \|A_x\|_2^2.$$

Also d/n is (asymptotically) the variance of the off-diagonal entries, for $x \neq y$,

$$\mathrm{Var}(A_{xy}) = \frac{d}{n}\left(1 - \frac{d}{n}\right).$$

Erdős-Rényi random graph

With $d' = d(1 - d/n)$, let

$$H = \frac{A - \mathbb{E}A}{\sqrt{d'}}$$

be the normalized adjacency of the Erdős-Rényi graph with average degree d.

$$\mathbb{E}A = \frac{d}{n}(J - I).$$

H is a sparse Wigner matrix with sparsity parameter

$$q = \frac{1}{\max |H_{xy}|} = \sqrt{d'}.$$

Perturbation inequalities

Let $A, B \in M_n(\mathbb{C})$ be Hermitian matrices.

Rank inequality, with $r = \operatorname{rank}(A - B)$ and for $i \geqslant 1$, $\lambda_{1-i} = \infty, \lambda_{n+i} = -\infty$.

$$\lambda_i(A) \leqslant \lambda_{i-r}(B).$$

Hoeffman-Wielandt inequality,

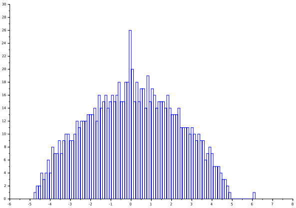
$$\sum_{i=1}^{n} (\lambda_i(A) - \lambda_i(B))^2 \leqslant \operatorname{Tr}(A - B)^2.$$

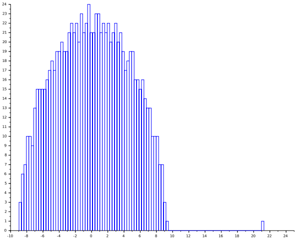
In the study of ESD's, these inequalities can be used to truncate and recenter variables. For the Erdős-Rényi graph

$$\left| \int \varphi d\mu_H - \int \varphi d\mu_{A/\sqrt{d'}} \right| \leqslant \int |\varphi|'/n + \|\varphi'\|_\infty \sqrt{d}/n.$$

Histogram of Eigenvalues of A

Single realization for $n = 1000$, $d = 5$ and $d = 20$.





Semicircle law for Wigner matrices

Let $H \in M_n(\mathbb{C})$ be an Hermitian random matrix with independent centered entries $(H_{xy})_{x \geqslant y}$ above the diagonal,

$$\text{for all } x, y, \quad \mathbb{E}|H_{xy}|^2 = \frac{1}{n} \quad \text{and} \quad a.s.\!-\max_{x,y} |H_{xy}| \leqslant \frac{1}{q}.$$

Theorem (Wigner's semicircle law)

If $q \to \infty$, *a.s.*

$$\lim_{n\to\infty} \mu_H = \mu_{sc}$$

where $d\mu_{sc}(\lambda) = \frac{\mathbf{1}(|\lambda| \leqslant 2)}{2\pi} \sqrt{4 - \lambda^2} d\lambda.$

The heart of the proof is the following lemma.

Lemma

If $q \to \infty$, for any integer k,

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}\mathrm{Tr}H^k = \int \lambda^k d\mu_{sc}(\lambda).$$

A direct computation gives

$$\int \lambda^{2k+1} d\mu_{sc}(\lambda) = 0 \quad \text{and} \quad \int \lambda^{2k} d\mu_{sc}(\lambda) = C_k,$$

where C_k is the k-th Catalan number:

$$C_k = \frac{1}{k+1}\binom{2k}{k} = \binom{2k}{k} - \binom{2k}{k+1}.$$

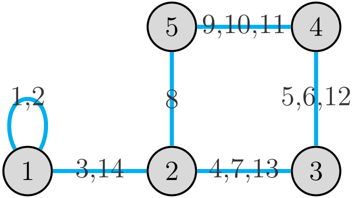
Expected trace method

We have

$$\frac{1}{n}\mathbb{E}\mathrm{Tr}H^k = \frac{1}{n}\sum_{\gamma} P(\gamma) \qquad \text{with} \quad P(\gamma) = \mathbb{E}\prod_{t=1}^{k} H_{\gamma_{t-1}\gamma_t},$$

where the sum is over all closed paths of length k,
$\gamma = (\gamma_0, \ldots, \gamma_k)$ with $\gamma_0 = \gamma_k$.

To each path, we associate a connected graph $G(\gamma)$ with $v(\gamma)$ vertices and $e(\gamma)$ edges.



$$\gamma = (1,1,1,2,3,4,3,2,5,4,5,4,3,2,1), \quad e(\gamma) = 6, \; v(\gamma) = 5.$$

Expected trace method

In $G(\gamma)$, each edge $e = (x, y), x \geqslant y$, is visited $m_e = m_e^+ + m_e^- \geqslant 1$ times. From the independence of entries

$$P(\gamma) = \mathbb{E} \prod_{t=1}^{k} H_{\gamma_{t-1}\gamma_t} = \prod_{e \in G(\gamma)} \mathbb{E}(H_e^{m_e^+} \bar{H}_e^{m_e^-}).$$

Since $\mathbb{E} H_{xy} = 0$: $P(\gamma) = 0$ unless for all e, $m_e \geqslant 2$. We thus consider only closed paths which visit each edge at least twice.

Since $\sum m_e = k$, $|H_{xy}| \leqslant 1/q$ and $\mathbb{E}|H_{xy}|^2 = 1/n$:

$$|P(\gamma)| \leqslant q^{-\sum_e (m_e - 2)} n^{-e(\gamma)}$$

If for all e, $m_e^+ = m_e^- = 1$ then $2e(\gamma) = k$ and $P(\gamma) = n^{-k/2}$.

Expected trace method

Let us say that two paths γ, γ' are equivalent if $\gamma' = \sigma \circ \gamma = (\sigma(\gamma_0), \ldots, \sigma(\gamma_k))$ for some permutation $\sigma \in S_n$.

An equivalence class is an unlabeled path. In the equivalence class of γ, there are

$$n(n-1)\cdots(n-v(\gamma)+1) \leqslant n^{v(\gamma)}$$

other paths.

Recall $|P(\gamma)| \leqslant q^{-(k-2e(\gamma))} n^{-e(\gamma)}$. If $\mathcal{W}_k$ is the set of unlabeled closed paths of length k which visit each edge at least twice:

$$\left|\frac{1}{n}\mathbb{E}\mathrm{Tr} H^k\right| = \frac{1}{n}\sum_{\gamma} |P(\gamma)| \leqslant \sum_{\gamma \in \mathcal{W}_k} n^{v(\gamma)-e(\gamma)-1} q^{-(k-2e(\gamma))}.$$

Expected trace method

$$\left|\frac{1}{n}\mathbb{E}\mathrm{Tr}H^k\right| = \frac{1}{n}\sum_{\gamma} |P(\gamma)| \leqslant \sum_{\gamma\in\mathcal{W}_k} n^{v(\gamma)-e(\gamma)-1} q^{-(k-2e(\gamma))}.$$

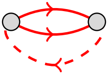
Since $G(\gamma)$ is connected, we have $v(\gamma) \leqslant e(\gamma)+1$ with equality if and only if $G(\gamma)$ is a tree.

$$\frac{1}{n}\mathbb{E}\mathrm{Tr}H^k = \sum_{\gamma\in\mathcal{W}_k^2} n^{k/2}P(\gamma) + o(1),$$

where $\mathcal{W}_k^2 \subset \mathcal{W}_k$ are the unlabeled paths such that $G(\gamma)$ is a tree and $k = 2e(\gamma)$.

In particular, if k is odd, we have $\frac{1}{n}\mathbb{E}\mathrm{Tr}H^k = o(1)$.

If $\gamma \in \mathcal{W}_k^2$ then for all e, $m_e^+ = m_e^- = 1$ as otherwise the path must contain a cycle:

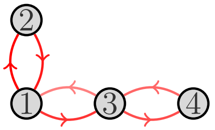


Hence, for $\gamma \in \mathcal{W}_k^2$, $P(\gamma) = n^{-k/2}$ and

$$\frac{1}{n}\mathbb{E}\mathrm{Tr} H^k = |\mathcal{W}_k^2| + o(1).$$

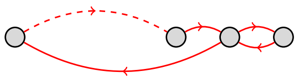
For $k \leqslant 2n-1$, $\mathcal{W}_k^2$ does not depend on n. It remains to check that $|\mathcal{W}_k^2| = C_{k/2}$. We use a bijective method.

We say that γ is canonical if $\gamma_0 = 1$ and $\gamma_t \leqslant \max_{s \leqslant t} \gamma_s + 1$ (i.e. the order of visits of the vertices is their value). There is a unique canonical path in each equivalence class.



$$\gamma = (1, 2, 1, 3, 4, 3, 1)$$

Let $\gamma \in \mathcal{W}_k^2$, either γ_t is visited for the first time or (γ_{t-1}, γ_t) is the last edge which has been visited once, as otherwise there is a cycle:



Expected trace method

To each canonical $\gamma \in \mathcal{W}_k^2$, we associate a path $x = (x_0, \dots, x_k)$ with $x_0 = 0$ and $x_{t+1} - x_t = +1$ if $\gamma_{t+1} = \max_{s \leqslant t} \gamma_s + 1$ and -1 otherwise.

$\gamma = (1, 2, 1, 3, 4, 3, 1)$ and its path x drawn as a function on $\{0, \dots, k\}$

This map is a bijection and its image is the set of Dyck paths

$$\mathcal{D}_{k/2} = \{(x_0, \dots, x_k) \in \mathbb{N}^{k+1} : x_0 = x_k = 0, x_{t-1} - x_t \in \{-1, 1\}\}.$$

In particular

$$|\mathcal{W}_k^2| = |\mathcal{D}_{k/2}|.$$

It is an exercise in combinatorics that $|\mathcal{D}_k| = C_k$. $\square$

Concentration

We have proved so far that $\mathbb{E}\mu_H$ converges to μ_{sc} weakly.

It remains to check that $\mathbb{E}\mu_H$ and μ_H are close with high probability.

First possibility: by similar combinatorial arguments, we may compute an upper bound on the variance:

$$\mathrm{Var}\left(\frac{1}{n}\mathrm{Tr} H^k\right) = O\left(\frac{1}{nq^2}\right).$$

Second possibility: general concentration inequalities apply here.

* (Azuma-Hoefding's inequality) If $\int |\varphi'| \leqslant 1$, for all $t \geqslant 0$,

$$\mathbb{P}\left(\left|\int \varphi d\mu_H - \mathbb{E}\int \varphi d\mu_H\right| \geqslant t\right) \leqslant 2\exp\left(-\frac{nt^2}{8}\right).$$

* (Talagrand's inequality) If $\|\varphi'\|_\infty \leqslant 1$ and has at most k inflections points, for all $t \geqslant 0$,

$$\mathbb{P}\left(\left|\int \varphi d\mu_H - \mathbb{E}\int \varphi d\mu_H\right| \geqslant t\right) \leqslant c_k\exp\left(-\frac{nq^2t^2}{c_k^2}\right).$$

Remarks

With $q \to \infty$, the same argument works inhomogeneous Wigner matrices with constant row variances:

$$\text{for all } x, \quad \mathbb{E}\sum_y |H_{xy}|^2 = 1 \quad \text{and} \quad a.s.{-}\max_{x,y} |H_{xy}| \leqslant \frac{1}{q}.$$

If $q = O(1)$ then the semicircle law may not hold. For example, for the Erdős-Rényi graph with fixed average degree d, the ESD converges to a probability measure which depends on d.

Some references

Wigner, E. (1955). *Characteristic vectors of bordered matrices with infinite dimensions.* Annals of Mathematics.

Anderson, G.W.; Guionnet, A.; Zeitouni, O. (2010). *An introduction to random matrices.* Cambridge: Cambridge University Press.

A. Bose, (2018) *Patterned random matrices.* CRC Press, Boca Raton.

Lecture notes on my webpage.

The Fűredi & Komlós upper bound

Convergence of the largest eigenvalue

Let $H \in M_n(\mathbb{C})$ be an Hermitian random matrix with independent centered entries $(H_{xy})_{x \geqslant y}$ above the diagonal,

$$\text{for all } x, y, \quad \mathbb{E}|H_{xy}|^2 = \frac{1}{n} \quad \text{and} \quad a.s.-\max_{x,y} |H_{xy}| \leqslant \frac{1}{q}.$$

The semi-circle law has support $[-2, 2]$. If $q \to \infty$, the Wigner's semicircle theorem implies that,

$$a.s.-\liminf_{n\to\infty} \lambda_1 \geqslant 2.$$

Theorem (Convergence of largest eigenvalue)

If $q \gg (\log n)^4$,

$$a.s.-\lim_{n\to\infty} \lambda_1 = 2.$$

This theorem is a direct consequence of:

Lemma (Fűredi & Komlós upper bound)

If $q \geqslant \sqrt{2}k^4$, *we have*

$$\mathbb{E}\mathrm{Tr}H^k \leqslant n2^{k+1}.$$

Expected high trace method

$$\mathbb{E}\mathrm{Tr}H^k \leqslant n \sum_{\gamma \in \mathcal{W}_k} n^{v(\gamma)-e(\gamma)-1} q^{-(k-2e(\gamma))},$$

where $\mathcal{W}_k$ is the set of unlabeled closed paths of length k which visit each edge at least twice.

Since $g(\gamma) = e(\gamma) + 1 - v(\gamma) \geqslant 0$ and $q^2 \leqslant n$, we have

$$\mathbb{E}\mathrm{Tr}H^k \leqslant n \sum_{\gamma \in \mathcal{W}_k} q^{-(k+2-2v(\gamma))} \left(\frac{q^2}{n}\right)^{g(\gamma)} \leqslant n \sum_{\gamma \in \mathcal{W}_k} q^{-(k+2-2v(\gamma))}.$$

Let $\mathcal{W}_k(v) \subset \mathcal{W}_k$ be the paths with $v(\gamma) = v$. If $v(\gamma) > \frac{k}{2} + 1$, $\mathcal{W}_k(v)$ is empty (because $e(\gamma) \leqslant k/2$ and $g(\gamma) \geqslant 0$).

$$\mathbb{E}\mathrm{Tr}H^k \leqslant n \sum_{v=1}^{k/2+1} |\mathcal{W}_k(v)| q^{-2(\frac{k}{2}+1-v)}.$$

Expected high trace method

$$\mathbb{E}\mathrm{Tr}H^k \leqslant n \sum_{v=1}^{k/2+1} |\mathcal{W}_k(v)| q^{-2(\frac{k}{2}+1-v)}.$$

Lemma

For any integer $1 \leqslant v \leqslant \frac{k}{2} + 1$,

$$|\mathcal{W}_k(v)| \leqslant 2^k k^{8(\frac{k}{2}+1-v)}.$$

We find, if $2k^8 \leqslant q^2$,

$$\mathbb{E}\mathrm{Tr}H^k \leqslant n2^k \sum_{v=1}^{k/2+1} \left(\frac{k^8}{q^2}\right)^{(\frac{k}{2}+1-v)} \leqslant n2^k \sum_{t\geqslant 0} \left(\frac{k^8}{q^2}\right)^t \leqslant n2^{k+1}.$$

$$|\mathcal{W}_k(v)| \leqslant 2^k k^{8(\frac{k}{2}+1-v)}.$$

We build an encoding of $\mathcal{W}_k(v)$, that is an injective map on $\mathcal{W}_k(v)$. Recall that unlabeled paths may be identified with canonical paths γ: $\gamma_0 = 1$ and $\gamma_t \leqslant \max_{s\leqslant t} \gamma_s + 1$ (i.e. the order of visits of the vertices is their value).

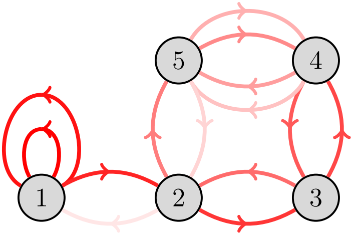
We mimic the Dyck path encoding of $\gamma \in \mathcal{W}_k^2$.

Let $\gamma = (\gamma_0, \ldots, \gamma_k) = (\gamma_t)_t$. Think of $t = 1, \ldots, k$ as a time. A time t is marked

- $+$ if γ_t is new (thus, the edge $\{\gamma_{t-1}, \gamma_t\}$ is seen for the first time);
- $-$ if the edge $\{\gamma_{t-1}, \gamma_t\}$ is seen for the second time and this edge was previously marked $+$;
- $\star$ otherwise.

Expected high trace method

$+$ new vertex, $-$ second visit of the edge, $\star$ otherwise.



$$\gamma = (1,1,1,2,3,4,3,2,5,4,5,4,3,2,1), \quad k = 14, \ v = 5$$

$$(\star, \star, +, +, +, -, -, +, \star, \star, \star, \star, -, -).$$

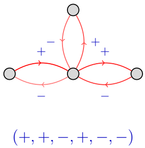
The number of $+$ is $v-1$. The number of $-$ is $v-1$. The number of $\star$ is $s = k - 2v + 2$.

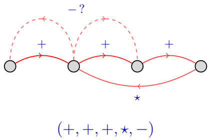
Expected high trace method

If the time is a $+$, then $\gamma_t = \max_{s<t} \gamma_s + 1$.

We mark the $\star$-times by their arrival vertex γ_t.

If the time is a $-$, there may be an ambiguity:





Call the ambiguous $-$ times, the crossroad times. We mark the crossroad times by their arrival vertex γ_t.

Fact: the number of crossroad times is at most the number of star times: $s = k - 2v + 2$.

The positions of $+$,$-$, $\star$ and *crossroad* times and their marks characterize uniquely a canonical path.

We get, with $s = k - 2v + 2$,

$$|\mathcal{W}_k(v)| \leqslant 2^k \cdot k^s \cdot k^s \cdot v^s \cdot v^s \leqslant 2^k k^{2s} v^{2s}.$$

Since $v \leqslant k$, we obtain the claim bound. $\square$

Remarks

The same argument works for inhomogeneous Wigner matrices with bounded row variances:

$$\text{for all } x, \quad \mathbb{E}\sum_y |H_{xy}|^2 \leqslant 1 \quad \text{and} \quad a.s.-\max_{x,y} |H_{xy}| \leqslant \frac{1}{q}.$$

Remarks

The condition $q \gg (\log n)^4$ for the convergence of λ_1 is not optimal.

For the Erdős-Rényi graph with average degree d and $H = (A - \mathbb{E}A)/\sqrt{d}$, we have $q = \sqrt{d}$ and

$$\lambda_1(H) = (1 + o(1))\frac{\lambda_2(A)}{\sqrt{d}} = 2 + o(1)$$

as soon as $q^2 = d \gg \log n$.

Remarks

If $(H_{xy})_{x \geqslant y}$ have a symmetric distribution and the matrix is flat, i.e. $q = O(\sqrt{n})$, then much better bound are available, A. Soshnikov (1999). O. Feldheim, S. Sodin (2010).

This allows to capture the fluctuations of the largest eigenvalue

$$\lambda_1 = 2 + (Z + o(1))n^{-2/3}.$$

Thanks to resolvent methods, the fluctuations are now known in a much greater generality: see J. Huang, B. Landon, H.T. Yau (2017) (if $q \geqslant n^{\alpha}, \alpha > 1/9$).

Remarks

General concentration inequalities apply here, for some universal constant c, for all $t \geqslant 0$,

$$\mathbb{P}(|\lambda_1 - \mathbb{E}\lambda_1| \geqslant t) \leqslant 2\exp(-cq^2t^2).$$

See the monographs,

S. Boucheron, G. Lugosi, P. Massart (2013) *Concentration Inequalities: A Nonasymptotic Theory of Independence*. Oxford University Press.

Anderson, G.W.; Guionnet, A.; Zeitouni, O. (2010). *An introduction to random matrices*. Cambridge: Cambridge University Press.

Further references

Fűredi, Z.; Komlós, J. *The eigenvalues of random symmetric matrices.* Combinatorica 1 (1981), no. 3, 233–241.

Vu, V.. *Spectral norm of random matrices.* Combinatorica 27 (2007), no. 6, 721–736.

Optimal bounds in the sparse regime: two recent works with F. Benaych-Georges, A. Knowles. See also, Latała, R; van Handel, R.; Youssef, P. *The dimension-free structure of nonhomogeneous random matrices.* Invent. Math. 214 (2018), no. 3, 1031–1080.

Bandeira & van Handel comparison argument

Comparison Argument

Let $H \in M_n(\mathbb{C})$ be an Hermitian random matrix with independent centered entries $(H_{xy})_{x \geqslant y}$ above the diagonal,

$$\text{for all } \ x, y, \quad \mathbb{E}|H_{xy}|^2 \leqslant \frac{1}{n} \quad \text{and} \quad a.s.{-}\max_{x,y} |H_{xy}| \leqslant \frac{1}{q}.$$

Let $S_r \in M_r(\mathbb{C})$ be a symmetric matrix with independent entries $(S_{xy})_{x \geqslant y}$ above the diagonal and with $p = 1/4$,

$$\text{for all } \ x, y, \quad S_{xy} \stackrel{d}{=} \frac{\mathrm{Ber}(p) - p}{\sqrt{p(1-p)}}.$$

Lemma

For all even k, *if* $r = \lceil q^2 \rceil + k$,

$$q^k \mathbb{E}\mathrm{Tr}(H^k) \leqslant \frac{n}{r} \mathbb{E}\mathrm{Tr}(S_r^k).$$

Comparison argument

We start by a lower bound on $\mathbb{E}\mathrm{Tr}(S_r^k)$.

$$\mathbb{E}\mathrm{Tr}(S_r^k) = \sum_{\gamma} Q(\gamma) \qquad \text{with} \qquad Q(\gamma) = \mathbb{E}\prod_{t=1}^{k} S_{\gamma_{t-1}\gamma_t}$$

where the sum is over all closed paths of length k, $\gamma = (\gamma_0, \ldots, \gamma_k)$ with $\gamma_0 = \gamma_k$.

Since the entries are iid above the diagonal and $\mathbb{E}S_{xy} = 0$

$$\mathbb{E}\mathrm{Tr}(S_r^k) = \sum_{\gamma \in \mathcal{W}_k} r(r-1)\cdots(r-v(\gamma)+1)Q(\gamma),$$

where $\mathcal{W}_k$ is the set of unlabeled closed paths of length k which visit each edge at least twice.

Comparison argument

Since $S_{xy} \overset{d}{=} \frac{\mathrm{Ber}(p)-p}{\sqrt{p(1-p)}}$ and $p = 1/4$, we may check that for all $m \geqslant 2$,

$$\mathbb{E}S_{xy}^m \geqslant 1.$$

It follows that $Q(\gamma) = \prod_{e \in G(\gamma)} \mathbb{E}(S_{xy}^{m_e}) \geqslant 1$ and

$$\mathbb{E}\mathrm{Tr}(S_r^k) \geqslant \sum_{\gamma \in \mathcal{W}_k} r(r-1)\cdots(r-v(\gamma)+1).$$

Set $r = \lceil q^2 \rceil + k$. Since $v(\gamma) \leqslant k$, we get $r - v(\gamma) + 1 \geqslant q^2$ and

$$\mathbb{E}\mathrm{Tr}(S_r^k) \geqslant \sum_{\gamma \in \mathcal{W}_k} q^{2v(\gamma)}.$$

If $r = \lceil q^2 \rceil + k$,

$$\mathbb{E}\mathrm{Tr}(S_r^k) \geqslant \sum_{\gamma \in \mathcal{W}_k} q^{2v(\gamma)}.$$

We have already proved the following upper bound on $\mathbb{E}\mathrm{Tr} H^k$.

$$\mathbb{E}\mathrm{Tr} H^k \leqslant n \sum_{\gamma \in \mathcal{W}_k} q^{-(k+2-2v(\gamma))}.$$

So that,

$$\mathbb{E}\mathrm{Tr} H^k \leqslant n q^{-k-2} \mathbb{E}\mathrm{Tr}(S_r^k).$$

As requested. $\square$

Application

The Fűredi-Komlós bound asserts that for $k = 4r^{1/8}$, $3/8 < \alpha < 1/2$,

$$\mathbb{E}\|S_r\| \leqslant (\mathbb{E}\mathrm{Tr} S_r^k)^{\frac{1}{k}} \leqslant (r^{k/2+1}2^{k+1})^{1/k} \leqslant 2\sqrt{r} + r^{\alpha}.$$

From Talagrand's inequality, for all k

$$\left(\mathbb{E}\mathrm{Tr} S_r^k\right)^{\frac{1}{k}} \leqslant r^{\frac{1}{k}} (\mathbb{E}\|S_r\|^k)^{\frac{1}{k}} \leqslant r^{\frac{1}{k}} (\mathbb{E}\|S_r\| + C\sqrt{k}).$$

If $r = \lceil q^2 \rceil + k$,

$$\mathbb{E}\lambda_1 \leqslant (\mathbb{E}\mathrm{Tr} H^k)^{\frac{1}{k}} \leqslant \left(\frac{n}{r}\right)^{\frac{1}{k}} q^{-1} \left(\mathbb{E}\mathrm{Tr} S_r^k\right)^{\frac{1}{k}}.$$

Optimizing over k, we find for some explicit $0 < \delta < 1$

$$\mathbb{E}\lambda_1 \leqslant 2 + C \max\left(\frac{\sqrt{\log n}}{q}, \left(\frac{\sqrt{\log n}}{q}\right)^{\delta}\right).$$

Convergence of the largest eigenvalue

Let $H \in M_n(\mathbb{C})$ be an Hermitian random matrix with independent centered entries $(H_{xy})_{x \geqslant y}$ above the diagonal,

$$\text{for all } x \neq y, \quad \mathbb{E}|H_{xy}|^2 = \frac{1}{n} \quad \text{and} \quad a.s.\!-\max_{x,y} |H_{xy}| \leqslant \frac{1}{q}.$$

Theorem (Convergence of largest eigenvalue)

If $q \gg \sqrt{\log n}$,

$$a.s.\!-\lim_{n\to\infty} \lambda_1 = 2.$$

Bordenave, Benaych-Georges & Knowles (2017)

Remarks

The same argument works for inhomogeneous Wigner matrices with bounded row variances:

$$\text{for all } x, \quad \mathbb{E}\sum_y |H_{xy}|^2 \leqslant 1 \quad \text{and} \quad a.s.-\max_{x,y}|H_{xy}| \leqslant \frac{1}{q}.$$

The bound $q \gg \sqrt{\log n}$ is optimal this time. For Erdős-Rényi graph it corresponds to average degree $d \gg \log n$.

Some references

A. Bandeira and R. van Handel. *Sharp nonasymptotic bounds on the norm of random matrices with independent entries*. Ann. Probab., 44(4):2479–2506, 2016.

Latała, R; van Handel, R.; Youssef, P. *The dimension-free structure of nonhomogeneous random matrices*. Invent. Math. 214 (2018), no. 3, 1031–1080.