Random matrix theory: lecture 17

Largest eigenvalue of Wigner's matrices

Preliminary: matrix norms

Def.: a norm on the space of $n\times n$ (complex) matrices is an application

$$\|\cdot\| : M_n \to \mathbb{R} \quad \text{such that}$$

- $\|A\| \geq 0$, $\|A\| = 0$ iff $A = 0$, $\forall A$
- $\|cA\| = |c| \cdot \|A\|$, $\forall c \in \mathbb{C}$, $\forall A$
- $\|A+B\| \leq \|A\| + \|B\|$, $\forall A, B$

Def: $\|\cdot\|$ is called a matrix norm if moreover

$$\|AB\| \leq \|A\| \cdot \|B\|, \quad \forall A, B$$

Properties: if $\|\cdot\|$ is a matrix norm, then

- $\|A^2\| \leq \|A\|^2$; likewise, $\|A^k\| \leq \|A\|^k$
- $A^2 = A \Rightarrow \|A\| = \|A^2\| \leq \|A\|^2 \Rightarrow \|A\| \geq 1$
- in particular, $\|I\| \geq 1$
- A invertible $\Rightarrow \|I\| = \|AA^{-1}\| \leq \|A\| \cdot \|A^{-1}\|$

$$\text{so } \|A^{-1}\| \geq \frac{\|I\|}{\|A\|} \geq \frac{1}{\|A\|}$$

Examples and counter-examples (⚠ notations ⚠)

1) The ℓ^1 norm $\|A\|_1 := \sum_{j,k=1}^{n} |a_{jk}|$ is a matrix norm.

2) The ℓ^2 norm (or Euclidean norm or Frobenius norm) defined as $\|A\|_2^2 := \sum_{j,k=1}^{n} |a_{jk}|^2$ is a matrix norm.

But note that the modified version of this norm:

$$\|A\|_{\tilde{2}}^2 := \frac{1}{n} \sum_{j,k=1}^{n} |a_{jk}|^2 \quad (= \|A\|_2^2 \text{ in homework 4})$$

is not a matrix norm ($\|AB\|_{\tilde{2}} \not\leq \|A\|_{\tilde{2}} \cdot \|B\|_{\tilde{2}}$ $\forall A, B$).

NB: both $\|\cdot\|_2$ and $\|\cdot\|_{\tilde{2}}$ are unitarily invariant, i.e.

$$\|UAV\|_2 = \|UA\|_2 = \|AV\|_2 = \|A\|_2 \quad \forall U, V \text{ unitary}$$

3) The ℓ^∞ norm $\|A\|_\infty := \max_{1 \leq j,k \leq n} |a_{jk}|$ is not a matrix norm but its modified version $\|A\|_{\tilde{\infty}} := n \max_{1 \leq j,k \leq n} |a_{jk}|$ is.

Induced norms

Def: let $\|\cdot\|$ be a (vector) norm on $\mathbb{C}^n$. We define the following induced norm on M_n:

$$|||A||| := \max_{\|x\|=1} \|Ax\| \quad \left(= \max_{x \neq 0} \frac{\|Ax\|}{\|x\|} \right)$$

Proposition: Induced norms are matrix norms.

In addition, $\|Ax\| \leq |||A||| \cdot \|x\|$ $\forall A, x$ and $|||I||| = 1$.

Examples:

1) The maximum column sum matrix norm $|||\cdot|||_1$ defined as

$$|||A|||_1 := \max_{1\le k\le n} \sum_{j=1}^{n} |a_{jk}|$$

is induced by the ℓ^1 norm on $\mathbb{C}^n$, i.e.

$$|||A|||_1 = \max_{\|x\|_1=1} \|Ax\|_1 \quad \text{where } \|x\|_1 = \sum_{j=1}^{n} |x_j| .$$

2) The spectral norm $|||\cdot|||_2$ defined as

$$|||A|||_2 := \max \{ \sqrt{\lambda} : \lambda \text{ is an eigenvalue of } A^*A \}$$

is induced by the ℓ^2 norm on $\mathbb{C}^n$, i.e.

$$|||A|||_2 = \max_{\|x\|_2=1} \|Ax\|_2 \quad (= \|A\|_1 \text{ in homework } 4)$$

NB: $|||\cdot|||_2$ is unitarily invariant.

3) The maximum row sum norm $|||\cdot|||_\infty$ defined as

$$|||A|||_\infty := \max_{1\le j\le n} \sum_{k=1}^{n} |a_{jk}|$$

is induced by the ℓ^∞ norm on $\mathbb{C}^n$, i.e.

$$|||A|||_\infty = \max_{\|x\|_\infty=1} \|Ax\|_\infty \quad \text{where } \|x\|_\infty = \max_{1\le j\le n} |x_j|$$

(Pfs: first show that $\|Ax\| \le |||A||| \cdot \|x\| \quad \forall x \in \mathbb{C}^n$
then find a x such that there is equality)

Spectral radius

$$\rho(A) := \max\{|\lambda| : \lambda \text{ is an eigenvalue of } A\}$$

Proposition

If A is normal, then $\rho(A) = \max\limits_{\|x\|_2=1} |x^* A x|$

(proof goes along the same lines as the proof of $|||A|||_2 = \max\limits_{\|x\|_2=1} \|Ax\|$)

Note that $\rho(\cdot)$ is <u>not</u> a norm, but the following holds:

Proposition

For any matrix norm $\|\cdot\|$, $\rho(A) \leq \|A\|$.

Proof

$\exists \lambda, x$ st $Ax = \lambda x$ and $|\lambda| = \rho(A)$.

Let X be the matrix whose columns are all equal to x.

Then $AX = \lambda X$ and for any matrix norm $\|\cdot\|$, we have

$$|\lambda| \cdot \|X\| = \|\lambda X\| = \|AX\| \leq \|A\| \cdot \|X\|$$

so $\rho(A) = |\lambda| \leq \|A\|$, since $\|X\| \neq 0$. #

Proposition (whose proof is more involved)

For any matrix norm $\|\cdot\|$, $\rho(A) = \lim\limits_{\ell\to\infty} \|A^\ell\|^{1/\ell}$.

[<u>ref</u>: Horn-Johnson p.299]

Back to Wigner's matrices

Let H^0 be a $n \times n$ real symmetric random matrix s.t.

(i) $\{h^0_{jk}, j \leq k\}$ are iid random variables (& $h_{kj} = h_{jk}$)

(ii) $|h^0_{11}(\omega)| \leq C \quad \forall \omega$ (bdd random variable)

(iii) $\mathbb{E}(h^0_{11}) = 0, \quad \mathbb{E}((h^0_{11})^2) = 1$

and $H^{(0,n)} := \frac{1}{\sqrt{n}} H^0$, $\lambda_1^{(0,n)} \ldots \lambda_n^{(0,n)} :=$ eigenvalues of $H^{(0,n)}$

We already know that

$$F_n^0(t) := \frac{1}{n} \# \{ j : \lambda_j^{(0,n)} \leq t \} \xrightarrow[n \to \infty]{} \int_{-\infty}^{t} p_\mu(x)\, dx \quad \text{a.s.}$$

where p_μ is the semi-circle distribution:

$$p_\mu(x) = \frac{1}{2\pi} \sqrt{4 - x^2} \cdot 1_{|x| \leq 2}$$

First remark

The same result holds if we replace assumption (iii) by the weaker assumption (iii)' $\mathbb{E}(h_{11}) = a \in \mathbb{R}$, $\mathrm{Var}(h_{11}) = 1$

Proof

× • Notice that $H = a\,\mathbb{1} + H^0$, where $\mathbb{1}$ is the all-ones matrix and H^0 satisfies assumptions (i)-(iii).

Therefore, $\mathrm{rank}(H - H^0) = 1$.

(NB: $a\mathbb{1}$ has 1 eigenvalue equal to na and $n-1$ zero eigenvalues)

- We will need <u>Weyl's inequalities</u>: if A, B are two real symmetric matrices with respective eigenvalues $\lambda_1^A \geq \ldots \geq \lambda_n^A$ and $\lambda_1^B \geq \ldots \geq \lambda_n^B$, then

$$\lambda_{j+k-1}^{A+B} \leq \lambda_j^A + \lambda_k^B \qquad \forall j,k \geq 1 \quad \text{s.t.} \quad j+k-1 \leq n$$

- Let $F_n(t) := \frac{1}{n} \# \{ j : \lambda_j^{(n)} \leq t \}$ with $\lambda_j^{(n)}$ = e.v. of $H^{(n)} = \frac{1}{\sqrt{n}} H$.

Then $\sup_{t \in \mathbb{R}} | F_n(t) - F_n^0(t) | \leq \frac{\text{rank}(H - H^0)}{n} = \frac{1}{n}$ $\quad (*)$

Therefore, F_n and F_n^0 converge to the same limit as $n \to \infty$. #

- Proof of $(*)$ using Weyl's inequalities:

Let $A = \frac{1}{\sqrt{n}} a \mathbb{1}$: since A is rank one, $\lambda_2^A = \ldots = \lambda_n^A = 0$

Let $B = H^{(0,n)}$; i.e. $\quad A + B = \frac{1}{\sqrt{n}} (a \mathbb{1} + H^0) = H^{(n)}$

Weyl's inequalities therefore imply that

$$\lambda_{k+1}^{(n)} \leq \lambda_k^{(0,n)} \qquad \forall k \leq n-1 \quad (\text{take } j=2)$$

Likewise, one can show that $\lambda_{k+1}^{(0,n)} \leq \lambda_k^{(n)}$.

So $|F_n(t) - F_n^0(t)| = \frac{1}{n} \left| \# \{ j : \lambda_j^{(n)} \leq t \} - \# \{ j : \lambda_j^{(0,n)} \leq t \} \right| \leq \frac{1}{n}$ #

<u>Conclusion</u>: The global regime (that of the semi-circle distribution) does not "see" the non-zero mean of the entries. The situation is quite different for local properties (such as the position of the largest eigenvalue).

Second remark

Wigner's result implies a lower bound on the largest eigenvalue of Wigner's matrices: for any $\varepsilon > 0$,

$$\frac{1}{n}\mathbb{E}\left(\#\{j : \lambda_j^{(n)} \in [2-\varepsilon, 2]\}\right) \xrightarrow[n\to\infty]{} \int_{2-\varepsilon}^{2} p_P(x)\,dx = c_\varepsilon > 0$$

i.e. the number of eigenvalues lying between $2-\varepsilon$ & 2 is equal to nc_ε in expectation. Therefore, there is in expectation at least one eigenvalue larger than $2-\varepsilon$,

i.e. $\lim_{n\to\infty} \mathbb{E}(\lambda_{max}(H^{(n)})) \geq 2-\varepsilon \quad \forall \varepsilon > 0$, i.e. ≥ 2.

Third remark

On the other hand, the largest eigenvalue of $H^{(n)}$ might be much greater than 2!

Example:

Consider $\mathbb{P}(h_{11} = 2) = \mathbb{P}(h_{11} = 0) = \frac{1}{2}$. Then $\mathbb{E}(h_{11}) = 1$, $Var(h_{11}) = 1$.
(satisfying hyp. (iii)')

It can easily be shown that

$$\sqrt{n} \leq \mathbb{E}\left(\lambda_{max}(H^{(n)})\right) \leq \sqrt{2n}$$

(→ homework)

Finally: let us show that in the case where

$$\mathbb{P}(h_{11} = +1) = \mathbb{P}(h_{11} = -1) = \frac{1}{2},$$

we have $\mathbb{E}(\lambda_{max}(H^{(n)})) \leq 2, \ \forall n.$

Proof:

- We will actually show that $\mathbb{E}(\rho(H^{(n)})) \leq 2, \ \forall n.$ implying that both the largest and lowest eigenvalues of $H^{(n)}$ lie in the interval $[-2, 2]$ in expectation.

- $\mathbb{E}(\rho(H^{(n)})) = \mathbb{E}\left(\lim_{\ell \to \infty} \|(H^{(n)})^\ell\|^{1/\ell}\right)$ for any matrix norm $\|\cdot\|$ (cf. preceding proposition page 4)

Let us choose $\|A\|^2 := \sum_{j,k=1}^n |a_{jk}|^2 = Tr(A^* A)$:

$$\mathbb{E}(\rho(H^{(n)})) = \mathbb{E}\left(\lim_{\ell \to \infty} Tr((H^{(n)})^{2\ell})^{1/2\ell}\right) \qquad ((H^{(n)})^* = H^{(n)})$$

$$\underset{\text{DCT}}{=} \lim_{\ell \to \infty} \mathbb{E}\left(Tr((H^{(n)})^{2\ell})^{1/2\ell}\right) \underset{\text{Jensen}}{\leq} \lim_{\ell \to \infty} \mathbb{E}\left(Tr((H^{(n)})^{2\ell})\right)^{1/2\ell}$$

- We have already computed $\mathbb{E}(Tr((H^{(n)})^{2\ell}))$: (see lect. 13)

$$= \frac{1}{n^\ell} \sum_{j_1 \cdots j_{2\ell} = 1}^n \mathbb{E}(h_{j_1 j_2} \cdots h_{j_{2\ell} j_1})$$

$$= \frac{1}{n^\ell} \sum_{g_{2\ell}} \sum_{J_{2\ell} \in g_{2\ell}} \mathbb{E}(h(J_{2\ell}))$$

$$= \frac{1}{n^\ell} \sum_{\substack{g_{2\ell} \text{ even} \\ |v(g_{2\ell})| \leq 1+\ell}} n(n-1) \cdots (n - |v(g_{2\ell})| + 1) \cdot \underbrace{Q(g_{2\ell})}_{\equiv 1 \text{ here, since } h_{jk}^{2m} \equiv 1 \ \forall m}$$

Note that allowing for "repetitions" of vertices, each graph with strictly less than $1+\ell$ vertices may be seen as a graph with $1+\ell$ vertices, allowing some of the vertices to be identical; and there are less than$^{(*)}$ $n^{1+\ell}$ such graphs, so

(*) since we are overcounting some graphs here

$$\mathbb{E}\left(\mathrm{Tr}\left(\left(H^{(n)}\right)^{2\ell}\right)\right) \leq \frac{1}{n^\ell} \sum_{\substack{g_{2\ell} \text{ even} \\ |v(g_{2\ell})| = 1+\ell}} n^{1+\ell}$$

Catalan numbers

$$= n \cdot \# \{ g_{2\ell} \text{ even} : |v(g_{2\ell})| = 1+\ell \} = n\, t_\ell$$

Therefore,

$$\mathbb{E}\left(\rho(H^{(n)})\right) \leq \lim_{\ell\to\infty} \left(n\, t_\ell\right)^{1/2\ell} = \lim_{\ell\to\infty} t_\ell^{\,1/2\ell} = 2, \quad \forall n \qquad \#$$

computed before (lect. 13)

($\underline{\text{NB}}$: $t_\ell = \int_{-2}^{2} x^{2\ell} \rho_{\Gamma}(x)\,dx$)

Remarks:

- The inequality $\rho(A) \leq \|A\|$ with the same choice of matrix norm $\|A\|^2 = \mathrm{Tr}(A^*A)$ does not suffice here:

$$\mathbb{E}\left(\rho(H^{(n)})\right) \leq \mathbb{E}\left(\sqrt{\mathrm{Tr}\left(\left(H^{(n)}\right)^2\right)}\right) \leq \sqrt{\frac{1}{n}\sum_{j,k=1}^{n} \underbrace{\mathbb{E}\left(h_{jk}^2\right)}_{=1}} = \sqrt{n}$$

and recall that $\|A\|^2 = \frac{1}{n}\mathrm{Tr}(A^*A)$ is $\underline{\text{not}}$ a matrix norm.

- Similarly, one can get bounds on $\mathbb{P}\left(\rho(H^{(n)}) \geq n^{\varepsilon}\right)$ using Chebyshev's inequality, but a more careful analysis is required for tighter bounds. [$\underline{\text{ref}}$: Soshnikov].