1 Catalan number

Recall that the Catalan number C_n equals

$$\frac{1}{n+1}\binom{2n}{n}.$$

If you want to know more different representations of Catalan numbers, you can read Richard Stanley's notes: `http://www-math.mit.edu/~rstan/bij.pdf`.

Note that we can use Talagrand's inequality to conclude that the mean or the median of $\|M\|$ is at least $(C_{k/2}^{1/k} + o_k(1))\sqrt{n}$. In particular, we would like to know the value of $C_{k/2}^{1/k}$ when k is large. By Stirling's formula, which says

$$k! \sim \sqrt{2\pi k}(k/e)^k,$$

we have

$$\begin{aligned} C_{k/2}^{1/k} &= \left(\frac{1}{k/2+1}\binom{k}{k/2}\right)^{1/k} \\ &\sim \left(\frac{\sqrt{2\pi k}(k/e)^k}{\pi k(k/2e)^k}\right)^{1/k} \\ &\sim 2. \end{aligned}$$

That is, $C_{k/2}^{1/k} \to 2$ as $k \to \infty$. This gives the mean or the median of $\|M\|$ is at least $(2-o(1))\sqrt{n}$ (also known as lower Bai-Yin Theorem). With some more work (which we will not show here), one can also obtain that the mean or the median of $\|M\|$ is at most $(2+o(1))\sqrt{n}$. In fact, one can even show that if the entries are bounded, then

$$\lim_{n\to\infty} \frac{\|M\|}{\sqrt{n}} = 2.$$

It is known that under suitable hypothesis, $\|M\|$ is concentrated in the range $[2\sqrt{n} - O(n^{-1/6}), 2\sqrt{n}+O(n^{-1/6})]$, and the normalized sequence $(\|M\| - 2\sqrt{n})n^{1/6}$ converges weakly to the Tracy-Widom distribution. I will talk about this in a separate notes, since it is interesting but we will not (be able to) cover this topic in the seminar.

This gives the heuristic that the eigenvalues of an $n \times n$ Wigner Hermitian matrix, after divided by $\sqrt{n}$, should concentrate on $[-2, 2]$. We will see that this is the case in certain sense.

2 The semicircle law

Definition 2.1. *Given any $n \times n$ Hermitian matrix M_n, the empirical spectral distribution (ESD), $\mu_{\frac{1}{\sqrt{n}}M_n}$, is defined to be the measure*

$$\mu_{\frac{1}{\sqrt{n}}M_n} = \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i(\frac{1}{\sqrt{n}}M_n)}.$$

We will show that the ESD of a Hermitian random matrix will converge in some senses, but first we have to define what convergence is for "random measures".

Definition 2.2. *Consider a sequence of Hermitian matrices M_n. We say*

$$\mu_{\frac{1}{\sqrt{n}}M_n} \to \mu \in \Pr(\mathbb{R})$$

in probability (resp. almost surely) if for all test function $\varphi \in C_c(\mathbb{R})$,

$$\int_{\mathbb{R}} \varphi \, \mathrm{d}\mu_{\frac{1}{\sqrt{n}}M_n} \to \int_{\mathbb{R}} \varphi \, \mathrm{d}\mu$$

in probability (resp. almost surely).

We will also define convergence in expectation.

Definition 2.3. *The expectation of $\mu_{\frac{1}{\sqrt{n}}M_n}$, denoted by $\mathbf{E}\mu_{\frac{1}{\sqrt{n}}M_n}$, is defined to be the probability measure such that*

$$\int_{\mathbb{R}} \varphi \, \mathrm{d}\mathbf{E}\mu_{\frac{1}{\sqrt{n}}M_n} = \mathbf{E} \int_{\mathbb{R}} \varphi \, \mathrm{d}\mu_{\frac{1}{\sqrt{n}}M_n}.$$

We say that the ESD's converge in expectation if

$$\int_{\mathbb{R}} \varphi \, \mathrm{d}\mathbf{E}\mu_{\frac{1}{\sqrt{n}}M_n} \to \int_{\mathbb{R}} \varphi \, \mathrm{d}\mu$$

for all test function φ.

Note that the expectation is always defined and is unique by Riesz representation theorem. We also remark that if the ESD's converge almost surely, then they converge in probability and in expectation (for the latter, it is because the test functions are always bounded, and so we can apply bounded convergence theorem).

Theorem 2.4 (The Wigner semicircle law)**.** *Let M_n be the top left $n \times n$ minor of an infinite Wigner matrix $(\xi_{i,j})$. Then the ESD's $\mu_{\frac{1}{\sqrt{n}}M_n}$ converges almost surely to the Wigner semicircular distribution*

$$\mathrm{d}\mu_{\mathrm{sc}} := \frac{1}{2\pi}(4 - |x|^2)_+^{1/2} \mathrm{d}x.$$

To prove this, there are 3 main reductions. We can assume that

1. the coefficients are bounded;
2. the diagonal entries vanish;
3. n ranges over a lacunary sequence: it suffices to show that the convergence holds over a subsequence n_m, where $n_m := \lfloor (1+\varepsilon)^m \rfloor$ for some $\varepsilon > 0$.

To see 1 and 2, we will make use of the Weilandt-Hoffman inequality: for $n \times n$ Hermitian matrices A and B,

$$\sum_{j=1}^{n} |\lambda_j(A+B) - \lambda_j(A)|^2 \leq \|B\|_F^2,$$

where the eigenvalues are ordered such that $\lambda_j \leq \lambda_{j+1}$, and $\|B\|_F^2 = \mathrm{tr}(B^2)$ is the Frobenius norm.

Lemma 2.5. *For all $n \times n$ Hermitian matrices A, B, for all $\lambda \in \mathbb{R}$, and for all $\varepsilon > 0$, we have*

$$\mu_{\frac{1}{\sqrt{n}}(A+B)}(-\infty, \lambda) \leq \mu_{\frac{1}{\sqrt{n}}A}(-\infty, \lambda + \varepsilon) + \frac{1}{\varepsilon^2 n^2}\|B\|_F^2$$

and

$$\mu_{\frac{1}{\sqrt{n}}(A+B)}(-\infty, \lambda) \geq \mu_{\frac{1}{\sqrt{n}}A}(-\infty, \lambda - \varepsilon) - \frac{1}{\varepsilon^2 n^2}\|B\|_F^2.$$

Proof. We just prove the first inequality. Suppose that $\lambda_i(A+B)$ is the largest eigenvalue of $A+B$ that is less than λn, and suppose that $\lambda_j(A)$ is the largest eigenvalue of A that is less than $(\lambda + \varepsilon)\sqrt{n}$. The first inequality then can be written as $i \leq j + \frac{1}{\varepsilon^2 n}\|B\|_F^2$.

If $i \leq j$, then we are done. If $i > j$, then for $j < \ell \leq i$, we have $|\lambda_\ell(A+B) - \lambda_\ell(A)| \geq \varepsilon n$, because $\lambda_\ell(A) \geq (\lambda + \varepsilon)\sqrt{n}$ but $\lambda_\ell(A+B) < \lambda\sqrt{n}$. Therefore,

$$\sum_{k=1}^{n} |\lambda_k(A+B) - \lambda_k(A)|^2 \geq \varepsilon^2 (i-j) n.$$

By the Weilandt-Hoffman inequality, we have

$$\varepsilon^2 (i-j) n \leq \|B\|_F^2.$$

Rearranging, we obtain the desired inequality. $\square$

References

[1] Tao, Terence. Topics in random matrix theory. Graduate Studies in Mathematics, 132. American Mathematical Society, Providence, RI, 2012. x+282 pp. ISBN: 978-0-8218-7430-1

[2] Thompson, Brady. Talk on 3/2.