Lecturer: Ramon van Handel Scribe: Xinyi Chen, Kiran Vodrahalli

# Contents



# 1 Last Time

**Theorem 1.1.** *Let $X, Y$ be random vectors in $\mathbb{R}^n$ with independent entries such that $\boldsymbol{E}[X_i] = \boldsymbol{E}[Y_i]$. $\boldsymbol{E}[X_i^2] = \boldsymbol{E}[Y_i^2]$. Let $f : \mathbb{R}^n \to \mathbb{R}$ be $C^3$. Then*

$$|\boldsymbol{E}[f(X)] - \boldsymbol{E}[f(Y)]| \leq \frac{1}{6}\sum_{i=1}^{n} \|\frac{\partial^3 f}{\partial X_i^3}\|_\infty \boldsymbol{E}[|X_i|^3 + |Y_i|^3] \tag{1}$$

**Example 1.2.** Let $X$ be an $n \times n$ symmetric matrix s.t. $\{X_{ij} : i \geq j\}$ are independent with $\mathbf{E}[X_{ij}] = 0, \mathbf{E}[X_{ij}^2] = 1$ (the so-called Wigner Matrix). We are interested in what the spectrum of $X$ looks like.

Let's first do a rough computation - how large is a typical eigenvalue?

$$\begin{aligned}
\mathbf{E}[\frac{1}{n}\sum_{i=1}^{n}\lambda_i(X)^2] &= \frac{1}{n}\mathbf{E}[\mathrm{Tr}(X^2)] \\
&= \frac{1}{n}\sum_{i,j}\mathbf{E}[X_{ij}^2] \\
&= \frac{1}{n} * n^2 * 1 = n
\end{aligned} \tag{2}$$

Thus a typical eigenvalue of $X$ has size $\approx \sqrt{n}$.

# 2 Wigner Semicircle Distribution

The object of our interest in this section is the empirical spectral distribution of $X/\sqrt{n}$, that is we define a probability measure on $\mathbb{R}$

$$\mu_n := \mathbf{E}[\frac{1}{n}\sum_{i=1}^{n}\delta_{\lambda_i(X/\sqrt{n})}] \tag{3}$$

where $\delta$ denotes a point mass at each eigenvalue. Therefore, $\mu_n([a, b])$ is the expected fraction of eigenvalues which lie in $[a, b]$.

**Theorem 2.1.** *(Wigner) Semicircle Distribution.*

$$\mu_n(dx) \to \mu_{SC}(dx) = \frac{1}{2\pi}\sqrt{4-x^2}\,\mathbf{1}_{|x|\le 2}dx \tag{4}$$

*where the convergence is weak. Here we assume bounded moments* $\boldsymbol{E}[X_{ij}] = 0, \boldsymbol{E}[X_{ij}^2] = 1, \boldsymbol{E}[|X_{ij}|^3] \le C$.

Now we have some mysteries to answer:

1. Why is the distribution semicircular?
2. Why does the distribution of $X_{ij}$ not play a role in the definition of $\mu_{SC}$?

To prove this theorem, we'll rely on two steps:

1. Step 1: Show that the limit does not depend on the distribution of $X_{ij}$ — this will follow from **universality**.
2. Step 2: Choose $X_{ij} \sim \mathcal{N}(0,1)$. We'll first prove the theorem in this special case using Gaussian tools. The Gaussian has such nice properties we refer to it as the "Gaussian paradise", with the properties of rotational invariance, easy integration by parts, and the properties of Brownian motion.

This two-step proof is not essential to proving the Wigner semicircle theorem, but otherwise, the proof is rather technical.

## 2.1 Tools

First we show $\mu_n \to \mu_{SC}$. We don't want to look at matrix functions because our function behaves unpleasantly over matrices. The usual idea would be to look at

$$\int e^{itu}\mu_n(du) \to \int e^{itu}\mu_{SC}(du) \tag{5}$$

Characteristic functions are nicer, but $e^{itu}$ is inconvenient to look at for matrices due to requiring analyticity.

We note that weak convergence iff character function convergence.

A more convenient parametrization of the distribution is to look at its Stieltjes transform.

**Definition 2.2.** Stieltjes transform of $\mu_n$

$$S_{\mu_n(z)} := \int \frac{1}{u-z}\mu_n(du) \to \int \frac{1}{u-z}\mu_{SC}(du) \tag{6}$$

with $z \in \mathcal{C} \setminus \mathbb{R}$. We note that $\frac{1}{u-z}$ is replaced by the algebraic function $(X - zI)^{-1}$ in the matrix setting. It turns out that the convergence of the Stieltjes transform is equivalent to weak converge of the measure $\mu_n \to \mu_{SC}$.

**Lemma 2.3.**

$$\int f(x)\mu(dx) = \lim_{\epsilon\to 0}\frac{1}{\pi}\int f(x)\mathrm{Im}(S_{\mu_n}(x+i\epsilon))dx \tag{7}$$

*for all* $f \in C_b(\mathbb{R})$.

*Proof.*

$$\text{Im}\left(\frac{1}{u-x-i\epsilon}\right) = \frac{\epsilon}{\epsilon^2+(u-x)^2} \tag{8}$$

$$S_\mu(x+i\epsilon) = \int \frac{1}{u-x-i\epsilon}\mu(du) \tag{9}$$

Then note that $p(u) := \frac{1}{\pi}\frac{\epsilon}{\epsilon^2+u^2}$ integrates to 1 by recognizing the Cauchy distribution with parameter $\epsilon$. Thus,

$$\frac{1}{\pi}\text{Im}(S_\mu(x+i\epsilon)) = \int p(x-u)\mu(du) \tag{10}$$

is the density of a random variable $X_\epsilon + Z$, where $X_\epsilon \sim \text{Cauchy}(\epsilon)$ and $Z \sim \mu$. with $X_\epsilon$ is independent of $\mu$.

Thus

$$\begin{aligned}\int f(x)\frac{1}{\pi}\text{Im}(S_{\mu_n}(x+i\epsilon))dx &= \mathbf{E}[f(X_\epsilon+Z)] \\ &\to_{\epsilon\to 0} \mathbf{E}[f(Z)] = \int f d\mu\end{aligned} \tag{11}$$

since $X_\epsilon \to 0$ as $\epsilon \to 0$. Essentially, after convolution our distribution looks wavy with a peak centered at each $\lambda_i$ where the width is $\approx \epsilon$. Taking the width $\epsilon \to 0$ turns this wavy distribution into a point mass distribution with peaks at each $\lambda_i$ — the definition of $\mu = \frac{1}{n}\sum_{k=1}^n \delta_{\lambda_k}$.

Therefore we have a one-to-one correspondance between $\mu$ and $S_{\mu_n}(Z)$, and the Stieltjes Transform contains all information of $\mu$. □

**Remark 2.4.** It is a fact that if $S_{\mu_n}(Z) \to S_\mu(Z)$, then $\mu_n \to \mu$.

## 2.2 Step 1

We write

$$f(X) := \frac{1}{n}\text{Tr}\left[\left(\frac{1}{\sqrt{n}}X - zI\right)^{-1}\right] = \frac{1}{n}\sum_{i=1}^n \frac{1}{\lambda_i(\frac{1}{\sqrt{n}}) - z} \tag{12}$$

So

$$\mathbf{E}[f(X)] = \int \frac{1}{u-z}\mu_n(du) \tag{13}$$

is the Stieljes transform of $\mu_n$.

We claim that for $X, Y$ $n \times n$ symmetric matricies with independent entries $\mathbf{E}[X_{ij}] = \mathbf{E}[Y_{ij}] = 0, \mathbf{E}[X_{ij}^2] = \mathbf{E}[Y_{ij}^2] = 1, \mathbf{E}[|X_{ij}|^3 + |Y_{ij}|^3] \le C$, we have

$$|S_{\mu_n(\frac{X}{\sqrt{n}})}(z) - S_{\mu_n(\frac{Y}{\sqrt{n}})}(z)| \lesssim \frac{C}{|z|^4\sqrt{n}} \to 0 \tag{14}$$

Let's first calculate the derivative of the inverse of a matrix with implicit differentiation:

$$\frac{d}{dt}M(t)M(t)^{-1} = 0 = \frac{dM(t)}{dt}M(t)^{-1} + M(t)\frac{dM(t)^{-1}}{dt}$$
$$\frac{dM(t)^{-1}}{dt} = -M(t)^{-1}\frac{dM(t)}{dt}M(t)^{-1} \tag{15}$$

We now proceed to calculate several derivatives of $f(X)$, for $i > j$. Let $E_{ij}$ denote the all-zeros matrix with the exception of a single 1 at position $(i, j)$. Note that $\frac{\partial X}{\partial X_{ij}} = E_{ij}$.

$$\frac{\partial f(X)}{\partial X_{ij}} = \frac{1}{n\sqrt{n}}\text{Tr}\left[(\frac{X}{\sqrt{n}} - zI)^{-1}E_{ij}(\frac{X}{\sqrt{n}} - zI)^{-1}\right] \tag{16}$$

by applying the inverse matrix derivative to the definition of $f$.

Then,

$$\frac{\partial^3 f(X)}{\partial X_{ij}^3} = \frac{6}{n^{5/2}}\text{Tr}\left[(\frac{X}{\sqrt{n}} - zI)^{-1}E_{ij}(\frac{X}{\sqrt{n}} - zI)^{-1}E_{ij}(\frac{X}{\sqrt{n}} - zI)^{-1}E_{ij}(\frac{X}{\sqrt{n}} - zI)^{-1}\right] \tag{17}$$

Note that $\|(\frac{X}{\sqrt{n}} - zI)^{-1}\|_{spectral} \leq \frac{1}{\text{Im}(z)}$, analagously to $|(u - z)^{-1}| \leq \frac{1}{\text{Im}(z)}$.

Then, we have

$$\left|\frac{\partial^3 f(X)}{\partial X_{ij}^3}\right| \leq \frac{1}{n^{5/2}}\frac{1}{\text{Im}(z)^4} \cdot C \tag{18}$$

which implies

$$|\mathbf{E}[f(X)] - \mathbf{E}[f(Y)]| \lesssim \sum_{i\geq j}\frac{1}{n^{5/2}}\frac{1}{\text{Im}(z)^4}\mathbf{E}[|X_{ij}|^3 + |Y_{ij}|^3]$$
$$\leq \frac{C}{\sqrt{n}}\frac{1}{\text{Im}(z)^4} \tag{19}$$

### 2.3 Step 2

Now we live in the Gaussian paradise. We take $X_{ij} \sim \mathcal{N}(0,1)$ and show that $S_{\mu_n}(z) \to S_{\mu_{SC}}(z)$, applying the result from Step 1 to generalize from this case.

First we have

$$\frac{1}{n}\text{Tr}\left[\frac{X}{\sqrt{n}}(\frac{X}{\sqrt{n}} - zI)^{-1}\right] = 1 + \text{Tr}\left[\frac{z}{n}(\frac{X}{\sqrt{n}} - zI)^{-1}\right]$$
$$\mathbf{E}[\frac{1}{n}\text{Tr}\left[\frac{X}{\sqrt{n}}(\frac{X}{\sqrt{n}} - zI)^{-1}\right]] = 1 + z\mathbf{E}[\frac{1}{n}\text{Tr}\left[(\frac{X}{\sqrt{n}} - zI)^{-1}\right]]$$
$$= 1 + zS_{\mu_n}(z) \tag{20}$$

Now recall that integration by parts for a Gaussian $g \sim \mathcal{N}(0,1)$ is given by $\mathbf{E}[gf(g)] + \mathbf{E}[f'(g)]$. Then, we have applying integration by parts

$$\mathbf{E}[X_{ij}\left[(\frac{X}{\sqrt{n}} - zI)^{-1}\right]_{ij}] - \frac{1}{\sqrt{n}} = \mathbf{E}[\frac{\partial}{\partial X_{ij}}\left[(\frac{X}{\sqrt{n}} - zI)^{-1}\right]_{ij}]$$
$$= \mathbf{E}[\left[(\frac{X}{\sqrt{n}} - zI)^{-1}E_{ij}(\frac{X}{\sqrt{n}} - zI)^{-1}\right]_{ij}]$$
$$= \mathbf{E}[\left[(\frac{X}{\sqrt{n}} - zI)^{-1}\right]_{ij}^2 + \left[(\frac{X}{\sqrt{n}} - zI)^{-1}\right]_{ii}\left[(\frac{X}{\sqrt{n}} - zI)^{-1}\right]_{jj}] \tag{21}$$

for $i \neq j$. Thus,

$$\begin{aligned}\frac{1}{n}\mathbf{E}[\mathrm{Tr}\left[\frac{X}{\sqrt{n}}(\frac{X}{\sqrt{n}} - zI)^{-1}\right]] = \frac{1}{n}\left[\frac{1}{n}\mathrm{Tr}\left[(\frac{X}{\sqrt{n}} - zI)^{-2}\right]\right] \\ + \mathbf{E}[\left(\frac{1}{n}\mathrm{Tr}\left[(\frac{X}{\sqrt{n}} - zI)^{-1}\right]\right)^2] \\ - \frac{1}{n}\mathbf{E}[\frac{1}{n}\sum_{i=1}^{n}\left[(\frac{X}{\sqrt{n}} - zI)^{-1}_{ii}\right]^2]\end{aligned} \tag{22}$$

We immediately note that the first and last terms are $\mathcal{O}(1)$. Then, in the middle term, we would like to get the square inside the expectation outside the expectation, so we can apply Gaussian Poincaré: Recall that variance is $\lesssim \frac{1}{\sqrt{n}}$log Poincaré. Therefore, we have bounded the difference between the Stieljes transforms appropriately. Now, we see

$$\begin{aligned}-S_{\mu_n}(z)^2 + \mathcal{O}(1) = 1 + zS_{\mu_n}(z) \\ -S_{\infty}(z)^2 = 1 + zS_{\infty}(z)\end{aligned} \tag{23}$$

Solving using the quadratic formula, we get

$$S_{\infty}(z) = \frac{-1 \pm \sqrt{z^2 - \varphi}}{2} \tag{24}$$

Since in a previous lemma we saw that taking $\epsilon \to 0$ yields convergence to $\mu_{SC}$,

$$\begin{aligned}\lim_{\epsilon \to 0}\frac{1}{\pi}\mathrm{Im}(S_{\infty}(x + i\epsilon)) = \lim_{\epsilon \to 0}\mathrm{Im}\left[\pm\frac{\sqrt{x^2 - \epsilon^2 + 2i\epsilon x - 4}}{2\pi}\right] \\ = \mathrm{Im}\left[\pm\frac{\sqrt{x^2 - 4}}{2\pi}\right]\end{aligned} \tag{25}$$

Notice that the imaginary part is gained with $x^2 \leq 4$, or $|x| \leq 2$. Thus, we can re-write the imaginary part as

$$\mu_{SC}(dx) = \frac{\sqrt{4 - x^2}}{2\pi} \cdot \mathbf{1}_{|x| \leq 2}dx \tag{26}$$

where we take the positive part, as desired.