Lecture 5:

21/11/2017

Wigner matrices:

$$X_N = \frac{1}{\sqrt{N}}\begin{pmatrix} Y_1 & Z_{1,2} & & \\ & Y_2 & & \\ & & \ddots & \\ & & & Y_N \end{pmatrix}$$

(Y_i), $(Z_{i,j})$ - IID, indep.

$\mathbb{E}(Y_1) = \mathbb{E}(Z_{1,2}) = 0$, $\mathbb{E} Z_{1,2}^2 = 1$

Gaussian Wigner: $Y_1 \sim N(0,\sigma^2)$

$Z_{1,2} \sim N(0,1)$

Eigenvalue: $\lambda_1^N \le \dots \le \lambda_N^N$

$$L_N = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i^N}, \quad \bar{L}_N = \mathbb{E} L_N$$

for any $f: \mathbb{R} \to \mathbb{R}$ suitably integrable ($[0,\infty)$-meas.)

$$\int f(x)\, d\bar{L}_N(x) = \langle \bar{L}_N, f\rangle = \mathbb{E}\langle L_N, f\rangle = \mathbb{E}\left(\int f(x)\, dL_N(x)\right)$$

Functions of sym or herm. matrices:

for H herm. & $f: \mathbb{R} \to \mathbb{R}$, let $f(H)$ be defined as follows:
It has the same eigenvectors as H & an eigenvalue λ becomes $f(\lambda)$. I.e., $H = UDU^*$ for a unitary U & diagonal $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{pmatrix}$

& $f(H) = U f(D) U^*$

$$f(D) = \begin{pmatrix} f(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & f(\lambda_N) \end{pmatrix}$$

Note that this is consistent with def for $f(x) = x^k$, $k \ge 1$ integer
& also $f = \frac{\text{Poly}}{\text{Poly}}$.

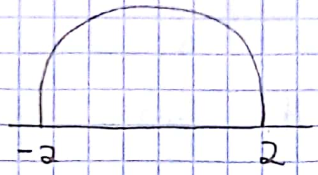
Concentration ineq.:

Suppose X_N has indep. entries $X_N(i,j)$, $j \ge i$ & is Herm,
Suppose that the distribution of each $X_N(i,j)$ satisfies the LSI with const. $\frac{c}{N}$ (e.g., Gaussian Wigner, $c = \max(\sigma^2, 1)$) & $f: \mathbb{R} \to \mathbb{R}$ is Lipschitz. Then:

$$P\left(\left|\underbrace{\frac{1}{N}\mathrm{tr}(f(X_N))}_{=\langle L_N, f\rangle} - \underbrace{\mathbb{E}\left[\frac{1}{N}\mathrm{tr}(f(X_N))\right]}_{\langle \bar{L}_N, f\rangle}\right| \ge \delta\right) \le 2e^{-\frac{N^2\delta^2}{4c\|f\|_{Lip}^2}}$$

Wigner's semicircle thm:

$$\sigma(x) = \frac{1}{2\pi}\sqrt{4-x^2}\,\mathbb{1}_{|x|\le 2}$$



Thm:

For any Wigner (with two finite moments) & $f:\mathbb{R}\to\mathbb{R}$ bdd. cont.,

$$\forall \varepsilon>0 \quad P(|\langle \mathcal{L}_N, f\rangle - \langle \sigma, f\rangle| > \varepsilon) \xrightarrow{N\to\infty} 0$$

(weak conv. in prob.)

Alternative approach to thm:

Stieltjes transform:

Let μ be a positive finite measure on $\mathbb{R}$. The Stieltjes transform of μ is $S_\mu(z) = \int \frac{\mu(dx)}{x-z} = \langle \mu, \frac{1}{x-z}\rangle$, $z \in \mathbb{C}\setminus\mathbb{R}$.

This converges as:

$$\left|\frac{1}{x-z}\right| = \frac{1}{|x-z|} \le \frac{1}{|\mathrm{Im}(z)|}$$

$$\frac{1}{N}\mathrm{Tr}((X_N - zI)^{-1}) = \langle \mathcal{L}_N, \frac{1}{x-z}\rangle$$

The Stieltjes transform of $\sigma(x)dx$ is

$$S_\sigma(z) = \frac{z}{2}\left(\sqrt{1-\frac{4}{z^2}} - 1\right) \qquad z \in \mathbb{C}\setminus[-2,2]$$

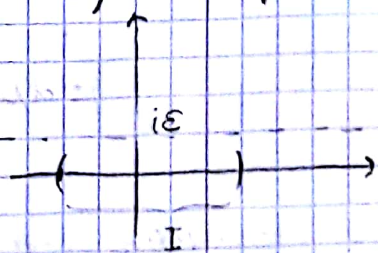
$$= \frac{\sqrt{z^2-4} - z}{2}$$

S_μ determines μ.

Thm:

For any open interval $I \subseteq \mathbb{R}$ with no atom of μ at endpoints,

$$\mu(I) = \lim_{\varepsilon\downarrow 0} \frac{1}{\pi}\int_I \mathrm{Im}(S_\mu(\lambda+i\varepsilon))\,d\lambda$$



$$\mathrm{Im}\left(\frac{1}{x-(\lambda+i\varepsilon)}\right) = \frac{\varepsilon}{(x-\lambda)^2+\varepsilon^2} > 0$$

Proof:

Note $\mu = 0$ iff $S_\mu \equiv 0$ since, e.g., $\mathrm{Im}(S_\mu(i)) = \int \frac{1}{x^2+1} \, d\mu(x) = 0$

$\Leftrightarrow \mu = 0$

Can recover $\int \mu(dx)$ since:

$y \,\mathrm{Im}(S_\mu(iy)) = \int \frac{y^2}{x^2+y^2} \, d\mu(x) \xrightarrow[y \to \infty]{} \int d\mu(x)$ bdd conv. thm.

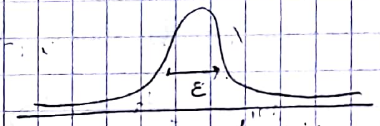
Assume WLOG, μ is a prob. measure.

Suppose X is a RV, $X \sim \mu$.

Observe that: $\frac{1}{\pi} \mathrm{Im}(S_\mu(\lambda + i\varepsilon)) = \frac{1}{\pi} \int \frac{\varepsilon}{(x-\lambda)^2 + \varepsilon^2} \, d\mu(x)$

which is the density at λ of $X + Y_\varepsilon$ where Y_ε is indep of X & has the cauchy dist.

density $(Y_\varepsilon) = \frac{\varepsilon}{\pi(x^2+\varepsilon^2)}$



density of Y_ε

Remains to note that $X + Y_\varepsilon \xrightarrow{d} X$ as $\varepsilon \downarrow 0$

(see, e.g. by Fourier transform $FT(Y_\varepsilon) \propto e^{-\varepsilon|t|}$)

This implies the thm.

Thm: (Relations with conv. of measures):

Let (μ_n) be a sequence of prob measures on $\mathbb{R}$.

1) If $\underbrace{\mu_n \xrightarrow{d} \mu}_{\int f d\mu_n \to \int f d\mu \ \forall \text{bdd. cont } f}$ then $S_{\mu_n}(z) \longrightarrow S_\mu(z) \quad \forall z \in \mathbb{C} \setminus \mathbb{R}$

(follows from definition)

2) Suppose $S_{\mu_n}(z) \longrightarrow S(z) \quad \forall z \in \mathbb{C} \setminus \mathbb{R}$, then $S(z)$ is the Stieltjes trans. of a sub-prob. meas. μ ($\int \mu(dx) \leq 1$) & $\underbrace{\mu_n \longrightarrow \mu \text{ Vaguely.}}_{\int f d\mu_n \longrightarrow \int f d\mu}$ for any $f: \mathbb{R} \to \mathbb{R}$, bdd. cont. & Vanishing at infinity.

Example: $\mu_n = \delta_n$

$\mu_n \longrightarrow 0$ Vaguely. $S_{\mu_n}(z) = \frac{1}{n-z} \longrightarrow 0 \quad \forall z \in \mathbb{C} \setminus \mathbb{R}$

3) Suppose μ_n are random & $S_{\mu_n}(z) \xrightarrow{\text{Prob.}} S_\mu(z)$ $\forall z \in \mathbb{C}\setminus\mathbb{R}$ where μ is a deterministic prob. meas.. Then:

$\mu_n \longrightarrow \mu$ in distribution in Prob.

$\left(P(|\langle\mu_n, f\rangle - \langle\mu, f\rangle| > \varepsilon) \xrightarrow{n\to\infty} 0 \text{ for any } f \text{ bdd. cont.} \right)$

Proof:

2) Helly's selection principle:

If μ_n is a seq. of Prob. meas. then

$\exists \mu$ sub-Prob. meas., n_k, with: $\mu_{n_k} \longrightarrow \mu$ Vaguely.

Vague convergence is metrizable. Thus enough that $\exists \mu$ sub.Prob.meas. for any n_k $\exists n_{k_l}$ with $\mu_{n_{k_l}} \longrightarrow \mu$ Vaguely.

Indeed, for each n_k $\exists n_{k_l}$ & μ s.t $\mu_{n_{k_l}} \longrightarrow \mu$ Vaguely.

whence $S_{\mu_{n_{k_l}}}(z) \longrightarrow \underset{S(z)}{\underset{\shortparallel}{S_\mu(z)}}$ $\forall z \in \mathbb{C}\setminus\mathbb{R}$ which means that μ is the same for all subsequences.

3) Idea (something in exercise):

Let d be a metric on prob. measures s.t. if $d(\mu_n, \mu) \xrightarrow{d} 0$ then $\mu_n \longrightarrow \mu$ in dist.

Then, if μ_n is a random seq. of prob. measures &

$P(d(\mu_n, \mu) > \varepsilon) \xrightarrow{n\to\infty} 0$ $\forall \varepsilon > 0$ Then $\mu_n \longrightarrow \mu$ weakly (in dist) in prob.

Let us build a suitable metric d. Let $z \in \mathbb{C}\setminus\mathbb{R}$, $(z_n) \subseteq \mathbb{C}\setminus\mathbb{R}$, $z_n \longrightarrow z$, $z_n \neq z$.

Define $d(\nu_k, \nu) = \sum_{i=1}^{\infty} 2^{-i} |S_{\nu_k}(z_i) - S_\nu(z_i)|$

$d(\nu_k, \nu) \longrightarrow 0$ iff $\forall n$ $S_{\nu_k}(z_n) \longrightarrow S_\nu(z_n)$

When this happens $\nu_k \longrightarrow \nu$ in dist. Since any subseq. of ν_k has a further subseq. conv. vaguely to some limit, but that limit must be ν, since its Stieltjes coincide with ν of (z_n) & thus everywhere by analyticity.

Remark: $S_\mu(z) = \int \frac{d\mu(x)}{x-z}$, $S_\mu(\bar{z}) = \overline{S_\mu(z)}$

Our assumption implies that $P(d(\mu_n,\mu)>\varepsilon) \xrightarrow{n\to\infty} 0$

<u>Wigner's thm via Stieltjes transform for Gaussian Wigner Matrices</u>

<u>Goal</u>: X_N Gaussian Wigner.

Show: $S_{Z_N}(z) \xrightarrow{\text{Prob}} S_\sigma(z) = \frac{z}{2}\left(\sqrt{1-\frac{4}{z^2}}-1\right) \quad \forall z\in\mathbb{C}\setminus\mathbb{R}.$

Convenient to start with $S_{\bar{Z}_N}(z) = \mathbb{E}\, S_{Z_N}(z)$.

<u>Two tricks</u>:

1) For an Herm. matrix X, define:

$$R_X(z) = S_X(z) := (X - zI)^{-1} \quad \text{(resolvent of } X)$$

<u>Identity</u>: $R_X(z) = \frac{1}{z}\left(XR_X(z) - I\right), \quad \forall z\in\mathbb{C}\setminus\mathbb{R}$

2) If $\xi \sim N(0,\sigma^2)$, RV & $f:\mathbb{R}\to\mathbb{R}$ cont. diff, of poly. growth of f, f', then:

$$\mathbb{E}(\xi f(\xi)) = \sigma^2\, \mathbb{E}(f'(\xi))$$

<u>Proof</u>:

$$\mathbb{E}(\xi f(\xi)) = \int x f(x)\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}dx = \int f'(x)\frac{\sigma^2}{\sqrt{2\pi}\sigma}e^{-x^2/2\sigma^2}dx$$

int. by parts

$f \to f'$

$$x\cdot\frac{1}{\sqrt{2\pi}\sigma}e^{\frac{-x^2}{2\sigma^2}} \longrightarrow \frac{-\sigma^2}{\sqrt{2\pi}\sigma}e^{-\frac{x^2}{2\sigma^2}}$$

<u>Calculation</u>:

What is the derivative of $R_X(z)$, with respect to X?

Real Symmetric case: $\frac{\partial}{\partial X}R_X(z) = -R_X(z)\Delta^{i,k}R_X(z)$ — means that we perturb X by adding ε to $X_{i,k}$ & $X_{k,i}$

$$\Delta^{i,k} = \begin{cases} \begin{pmatrix} 0&0&1&0&\cdots\\ 1&0&0&\cdots \\ \vdots \end{pmatrix} & \text{pos}(i,k),\ \ell(k,i) \quad i\neq k \\ \begin{pmatrix} 0&0&0\\ 0&1&0\\ 0&0&0\end{pmatrix} & \text{pos}(i,i) \quad i=k \end{cases}$$

Indeed, if E is hermitian, then $R_{X+\varepsilon E}(z) = (X+\varepsilon E - zI)^{-1} =$

$$= \left[(X-zI)\left(I + (X-zI)^{-1}\varepsilon E\right)\right]^{-1} = \underbrace{\left(I+\varepsilon(X-zI)^{-1}E\right)^{-1}}_{I-\varepsilon(X-zI)^{-1}E+\varepsilon^2((X-zI)^{-1}E)^2+\cdots}R_X(z)$$

$$= R_X(z) - \varepsilon R_X(z)ER_X(z) + O(\varepsilon^2)$$

$$S_{\bar{Z}_N}(z) = \langle \bar{Z}_N, \frac{1}{x-z} \rangle = \mathbb{E} \langle Z_N, \frac{1}{x-z} \rangle = \frac{1}{N} \mathbb{E}\, tr\, [R_{X_N}(z)] =$$

tricke

$$= \frac{1}{N} \mathbb{E}\, tr\, \frac{1}{z}(X_N R_{X_N}(z) - I) = -\frac{1}{z} + \frac{1}{Nz} \underbrace{\mathbb{E}\, tr(X_N R_{X_N}(z))}_{\text{use trick 2}}$$

Write $X_N = \sum_{\substack{i,k=1 \\ k \geq i}}^{N} X_N(i,k) \Delta^{i,k}$

$$\sigma^2 = \begin{cases} \frac{1}{N} & i \neq k \\ \frac{\mathbb{E}(Y_i^2)}{N} & i = k \end{cases}$$

$$X_N R_{X_N}(z) = \sum_{\substack{i,k=1 \\ k \geq 1}}^{N} X_N(i,k) \Delta^{i,k} R_{X_N}(z) \overset{\text{trick (2)}}{=} \sigma^{-2} \mathbb{E}(f')$$

$$\mathbb{E}\, tr(X_N R_{X_N}(z)) = \sum_{\substack{i,k=1 \\ k \geq i}}^{N} \mathbb{E}\Big[X_N(i,k)\, \underbrace{tr(\Delta^{i,k} R_{X_N}(z))}_{= f(X_N(i,k)) \text{ with all other entries of } X_N \text{ fixed.}}\Big]$$

<u>Conclude</u>:

$$S_{\bar{Z}_N}(z) = \frac{1}{z} - \frac{1}{N^2 z} \mathbb{E}\Big[\sum_{i,k=1}^{N} R_{X_N}(z)(i,i) R_{X_N}(z)(k,k) + R_{X_N}(i,k)^2\Big] -$$

$$- \frac{1}{N^2 z} \sum_i (\mathbb{E}(Y_i^2) - 2)\, \mathbb{E}(R_{X_N}(i,i)^2)$$

$$\sum_{i,k} R_{X_N}(z)(i,i) R_{X_N}(z)(k,k) = \Big(\sum_i R_{X_N}(i,i)\Big)^2 = (tr(R_{X_N}))^2 =$$

$$= \Big(\langle Z_N, \frac{1}{x-z} \rangle N\Big)^2$$

$$\sum_{i,k} R_{X_N}(z)(i,k)^2 = tr(R_{X_N}(z) \underbrace{R^T_{X_N}(z)}_{= R_{X_N}(z)}) = tr(R^2_{X_N}(z)) = \langle Z_N, \frac{1}{(x-z)^2} \rangle N$$

Thus,

can put $\bar{Z}_N$ by concentration.

$$S_{\bar{Z}_N}(z) = -\frac{1}{z} - \frac{1}{z} \mathbb{E}(\langle Z_N, \frac{1}{x-z} \rangle^2) - \underbrace{\frac{1}{Nz} \langle \bar{Z}_N, \frac{1}{(x-z)^2} \rangle}_{o(1)} - \underbrace{\frac{1}{N^2 z} \sum_i (\mathbb{E}(Y_i^2) - 2)\mathbb{E}(R^2_{X_N}(i,i))}_{= o(1)}$$

$$\langle \bar{Z}_N, \frac{1}{x-z} \rangle$$

since $\frac{1}{x-z}$ is bdd in x when $z \notin \mathbb{R}$

since that $\frac{1}{N^2} \sum_{i,k} \mathbb{E}(R_N(i,i)^2) = o$

$$S_{\bar{Z}_N}(z) = -\frac{1}{z} - \frac{S_{\bar{Z}_N}(z)^2}{z} + o(1) \quad \text{as } N \to \infty$$

We conclude that:

$$S_{Z_N}(z) = -\frac{1}{z} - \frac{S_{\bar{Z}_N}(z)^2}{z} + o(1) \quad \text{as } N \to \infty.$$

The limit equation $S(z) = -\frac{1}{z} - \frac{S^2(z)}{z}$ which means

$$S(z) = \frac{z}{2}\Big(\pm\sqrt{1 - \frac{4}{z^2}} - 1\Big)$$

We have the + sign since $\mathrm{Im}(S_{\bar{Z}_N}(\lambda + i\varepsilon)) > 0$, $\varepsilon > 0$.

We deduce that $S_{\bar{Z}_N}(z) \longrightarrow \frac{z}{2}\left(\sqrt{1 - \frac{4}{z^2}} - 1\right)$ for any $z \in (\mathbb{C} \setminus \mathbb{R})$.

This implies that $S_{Z_N}(z) \xrightarrow{P} \frac{z}{2}\left(\sqrt{1 - \frac{4}{z^2}} - 1\right)$ since $S_{Z_N}(z) = \langle Z_N, \frac{1}{x-z} \rangle$

is close in prob. to $S_{\bar{Z}_N}(z) = \langle \bar{Z}_N, \frac{1}{x-z} \rangle$ by concentration ineq.