Lecture 4: 14/11/2017

Reminder:

Wigner matrices:

$$X_N = \frac{1}{\sqrt{N}} \begin{pmatrix} Y_1 & Z_{12} & & \\ Z_{21} & Y_2 & & \\ \vdots & & \ddots & \\ & & & Y_N \end{pmatrix}$$

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

$\forall k \geq 1 \quad r_k = \max\left(\mathbb{E}(|Y_1|^k), \mathbb{E}(|Z_{1,2}|^k)\right) < \infty$

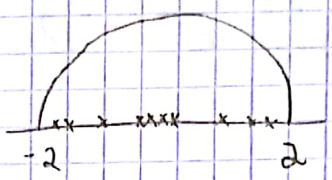
(Y_i) IID, (Z_{ij}) IID

independent.

Eigenvalues:

$\lambda_1^N \leq \dots \leq \lambda_N^N$, $\quad \mathcal{L}_N = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i^N}$

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



Thm (Wigner):

$\forall f \in C_b(\mathbb{R}) \qquad P\left(\left| \underbrace{\langle \mathcal{L}_N, f\rangle}_{=\frac{1}{N}\sum_{i=1}^N f(\lambda_i^N)} - \langle \sigma, f\rangle \right| > \varepsilon\right) \xrightarrow{N\to\infty} 0$

Maximal eigenvalue:

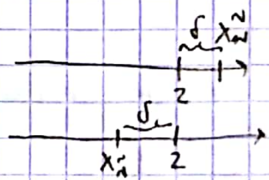
Thm.:

Suppose that $\exists C, c > 0$ s.t $r_k \leq (Ck)^{ck}$ for all $k \geq 1$

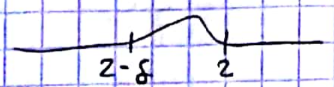
(E.g. $\exists C, \alpha > 0 \quad P(|Z_{1,2}| > t) \leq Ce^{-\alpha t}$, $P(|Y_1| > t) \leq Ce^{-\alpha t}$)

then $\lambda_N^N \longrightarrow 2$ in Probability.

Proof:

Idea: we want to show that  is unlikely & is unlikely

First, show λ_N^N is not too small. Indeed, let $\delta > 0$, take $f : \mathbb{R} \to \mathbb{R}$, $\int f\sigma(x)dx > 0$, cont., bdd., supported on $[2-\delta, 2]$



$P(\lambda_N^N < 2-\delta) \leq P\left(\int f(x)\, d\mathcal{L}_N(x) = 0\right) \leq P\left(\left|\int f(x)\, d\mathcal{L}_N(x) - \int f(x)\sigma(x)dx\right| > \frac{1}{2}\int f(x)\sigma(x)dx\right)$

Wigner $\downarrow N\to\infty$ 0

Intuition:

Observe that $\int x^{2k}\, dL_N(x) = \frac{1}{N}\sum_{i=1}^{N} (\lambda_i^N)^{2k} \geq \frac{1}{N}(\lambda_N^N)^{2k}$

$\int x^{2k}\, dL_N(x) \xrightarrow[N\to\infty]{} C_k \sim 4^k k^d$

"in limit"

$\Rightarrow \lambda_N^N \leq (N C_k)^{\frac{1}{2k}}$, here N isn't necessarily neglible

The idea: we will take an extremely large $k = (\omega \log N)$

Proof:

Fix $\delta > 0$. $P(\lambda_N^N > 2+\delta) = P\left(\int x^{2k}\, dL_N(x) > \frac{1}{N}(2+\delta)^{2k}\right) \leq$

$\leq \frac{N\,\mathbb{E}\left(\int x^{2k}\, dL_N(x)\right)}{(2+\delta)^{2k}} \leq$ we need to bound the numerator.

Reminder:

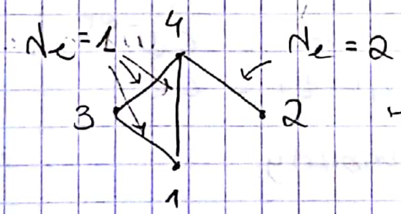
$$\mathbb{E}\left(\int x^{2k}\, dL_N(x)\right) = \sum_{t=1}^{k+1} \frac{C_{N,t}}{N^{k+1}} \sum_{\omega \in W_{2k,t}} \underbrace{\prod_{e \in E_\omega^c} \mathbb{E}\left(Z_{1,2}^{N_e^\omega}\right) \prod_{e \in E_\omega^s} \mathbb{E}\left(Y_1^{N_e^\omega}\right)}_{(*)}$$

$C_{N,t} = N(N-1)\cdot\ldots\cdot(N-t+1) \leq N^t$

$W_{2k,t}$ = Number of equiv. classes of words $i_1, i_2, \ldots, i_{2k}, i_1$ each $i_j \in \{1,\ldots,N\}$ having exactly t numbers & $N_e^\omega \geq 2$ for each e in this graph.

Example:

$1\,3\,4\,2\,4\,1$



$\leftarrow$ graph(V_ω, E_ω)

Know: $W_{2k,k+1} = C_k \longrightarrow$ these are the trees. the catalan number counts the rooted trees.

Lemma: (Combinatorics - We will skip it)

For all $t \leq \frac{k+1}{2}$: $|W_{k,t}| \leq 2^k k^{3(k-2t+2)}$

Using this:

$$\mathbb{E}\left(\int x^{2k}\, dL_N(x)\right) \leq \sum_{t=1}^{k+1} \frac{1}{N^{k+1-t}} |W_{2k,t}| \max_{\omega \in W_{2k,t}} (*) \leq$$

$$\leq 4^k \sum_{t=1}^{k+1} \left(\frac{(2k)^6}{N}\right)^{(k+1-t)} \max_{\omega \in W_{2k,t}} (*)$$

Suppose $\omega \in W_{2k,t}$ & $\#\{e \in E_\omega^c \mid N_e^\omega = 2\} = l$, then:

$$\prod_{e \in E_\omega^c} \mathbb{E}(Z_{1,2}^{N_e^\omega}) \prod_{e \in E_\omega^s} \mathbb{E}(|Y_1|^{N_e^\omega}) \le r_{2k-2l}$$

(For a non negative real variable Z, $\mathbb{E}\, Z^m \le (\mathbb{E}\, Z^{m+n})^{\frac{m}{m+n}}$, $l \ge 0$
Holder or Jensen. & we will take $m+n = 2k-2l$)

Also, every e with $N_e^\omega \neq 2$ has $N_e^\omega \ge 3$. Since G_ω is connected, $t = |V_\omega| \le |E_\omega| + 1$. but $2k = 2l + 3(|E_\omega^c| - l) + 2|E_\omega^s|$

$\Longrightarrow l \ge 3(t-1) - 2k \Longrightarrow 2k - 2l \le 6(k+1-t)$

So $\max_{\omega \in W_{2k,t}} (*) \le r_{6(k+1-t)}$

Since r_m is monotone in m.

Thus, $\mathbb{E}\left(\int x^{2k} \, d\lambda_n(x)\right) \le \sum_{t=1}^{k+1} \left(\frac{(2k)^6}{N}\right)^{k+1-t} r_{6(k+1-t)} \le$

$\le 4^k \sum_{t=1}^{k+1} \left(\frac{(2k)^6 \, (ck)^6}{N}\right)^{k+1-t} \le 4^k \sum_{i=0}^{k} \left(\frac{k^\alpha}{N}\right)^i \le 4^k \frac{1}{1 - \frac{k^\alpha}{N}}$

$\frac{k^\alpha}{N} < 1$

$$\stackrel{(\le)}{} \quad \frac{N \cdot 4^k \cdot \frac{1}{1 - \frac{k^\alpha}{N}}}{(2+\delta)^{2k}} = \underbrace{\left(\frac{4}{2+\delta}\right)^k}_{\approx e^{-\delta k}} N \cdot \frac{1}{1 - \frac{k^\alpha}{N}} \xrightarrow{N \to \infty} 0$$

(if we take $k = N^{\frac{1}{\alpha}}$,
$\delta = \frac{c \log(N)}{N^{1/\alpha}}$, $\alpha = \frac{1}{6}$
then we will get the limit)

when we take $\frac{k^\alpha}{N} < 1$ and $\frac{k}{\log N} \longrightarrow \infty$

■

2.3 Logarithmic Sobolev inequalities:

Idea: $X_1, \ldots, X_n$ indep., $f: \mathbb{R} \to \mathbb{R}$ Lipshitz, then:
$f(X_1, \ldots, X_n)$ is close to $\mathbb{E}[f(X_1, \ldots, X_n)]$ with high prob.

Another Idea: Sub-Gaussian concentration:

$$P\left(\left|f(X_1, \ldots, X_n) - \mathbb{E}[f(X_1, \ldots, X_n)]\right| \ge t\right) \le c e^{-ct^2} \quad (*)$$

Whether it holds depends on the joint dist. of $(X_1, \ldots, X_n)$.

We want a cond. on each X_i separately to ensure $(*)$.

Definition:

$F: \mathbb{R}^m \to \mathbb{R}$ is called Lipschitz if $\sup\limits_{\substack{x,y\in\mathbb{R}^m \\ x\neq y}} \frac{|F(x)-F(y)|}{\|x-y\|_2} =: |F|_L < \infty$

Definition: (think of Prob. measures as the joint dist. of random variables)

A probability measure P on $\mathbb{R}^m$ satisfies the Log-Sobolev inequality (LSI) with constant $c>0$ if for every differentiable (cont.?) $F: \mathbb{R}^m \to \mathbb{R}$,

$$\int F^2 \log\left(\frac{F^2}{\int F^2 dP}\right) dP \leq 2c \int \|\nabla f\|_2^2 \, dP$$

Intuition: The growth of the averages of functions is bounded by the change in the ∇f.

Lemma 2.3.2: (not proved here).

Suppose P has a density of the form $e^{-V(x)}dx$ for some $V: \mathbb{R}^m \to \mathbb{R}\cup\{\infty\}$ satisfying that:

$V(x) - \frac{\|x\|_2^2}{2c}$ is convex.

Then P satisfies the LSI with constant c.

(In particular, the standard Gaussian satisfies LSI with const. 1)

Exercise: if $X_1, \dots, X_M$ are random var., the dist. of each of which satisfies LSI with constant c. Then if the (X_i) are indep. then the dist. of $(X_1, \dots, X_M)$ satisfies LSI with cons. c.

Lemma (Herbst):

Suppose P satisfies LSI with const. c. Suppose G is Lipshitz with const. $|G|_L$. Then for all $\lambda \in \mathbb{R}$, $X = (X_1, \dots, X_M) \sim P$

$$\mathbb{E}\left[e^{\lambda(G(x) - \mathbb{E}(G(x)))}\right] \leq e^{c\lambda^2 |G|_L^2/2} \quad (*)$$

Consequently, for $\delta > 0$,

$$P(|G(x) - EG(x)| \geq \delta) \leq 2e^{-\delta^2/2c|G|_2^2} \quad (**)$$

(we are also claiming that $EG(x)$ is well defined & finite)

Proof:

To get (**):

$$P(G(x) - EG(x) \geq \delta) = P(e^{\lambda G(x) - EG(x)} \geq e^{\lambda\delta}) \overset{\text{markov}}{\leq}$$

$$\leq \frac{E[e^{\lambda G(x) - EG(x)}]}{e^{\lambda\delta}}$$ & use (*) & optimize λ.

Case I: Assume G is bdd & continuously diff.

$$|G|_2^2 = \| \|\nabla G\|_2^2 \|_\infty = \sup_{x \in \mathbb{R}^m} \|\nabla G(x)\|_2^2$$

Take $F = e^{\lambda(G - EG)}$. Plugging in LSI:

$$E\, e^{2\lambda(G-EG)} \log\left[\frac{e^{2\lambda(G-EG)}}{E(e^{2\lambda(G-EG)})}\right] \leq 2c\lambda^2 E[e^{2\lambda(G-EG)} \|\nabla G\|_2^2] \leq$$

$$\leq 2c\lambda^2 |G|_2^2 E[e^{2\lambda(G-EG)}]$$

Rearranging,

$$\frac{d}{d\lambda}\left[\frac{A_\lambda}{\lambda}\right] \leq 2c|G|_2^2$$

where $A_\lambda = \log(E[e^{2\lambda(G-EG)}])$

So: $$\frac{A_t}{t} - \underbrace{\lim_{t \downarrow 0} \frac{A_t}{t}}_{\text{show } = 0} = \int_0^t \frac{d}{d\lambda}\left[\frac{A_\lambda}{\lambda}\right] d\lambda \leq 2c|G|_2^2 t$$

$\Rightarrow A_t \leq 2c|G|_2^2 t^2$ which is what we want with $t = \frac{\lambda}{2}$.

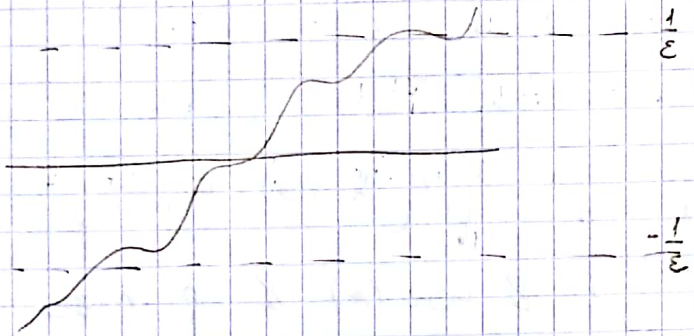
why is $\lim_{\lambda \downarrow 0} \frac{A_\lambda}{\lambda} = 0$?

G is bdd, so for small λ, have $e^{2\lambda(G-EG)} = 1 + 2\lambda(G - EG) + O(\lambda^2)$ as $\lambda \downarrow 0$

Case II : G is a general Lipschitz function.

For each $\varepsilon > 0$, define:

$$\bar{G}_\varepsilon(x) = \min\left(\max\left(G(x), -\tfrac{1}{\varepsilon}\right), \tfrac{1}{\varepsilon}\right)$$



$|\bar{G}_\varepsilon|_L \le |G|_L$ for all ε.

Take $P_\varepsilon(x) = \dfrac{1}{\sqrt{(2\pi\varepsilon)^m}} e^{-\frac{\|x\|^2}{2\varepsilon}}$ to be the density of

$(Y_1, \dots, Y_m)$ IID $N(0, \varepsilon)$ Random Vars

$G_\varepsilon = \bar{G}_\varepsilon * P_\varepsilon \iff G_\varepsilon(x) = \int \bar{G}_\varepsilon(y) P_\varepsilon(x-y)\,dy = \int \bar{G}_\varepsilon(x-y) P_\varepsilon(y)\,dy$

Thus, G_ε is smooth. (G_ε inherits the properties of P_ε as a function of x)

Also, $|G_\varepsilon|_L \le |\bar{G}_\varepsilon|_L \le |G|_L$ for all ε.

Applying Case I: for all $\lambda > 0$,

$$\mathbb{E}\left[e^{\lambda(G_\varepsilon(x) - \mathbb{E}(G_\varepsilon(x)))}\right] \le e^{c\lambda^2 |G|_L^2/2}. \quad (*)$$

To go back from G_ε to G:

First, $\forall x$, $G_\varepsilon(x) \xrightarrow{\varepsilon \to 0} G(x)$. Indeed,

$$|G_\varepsilon(x) - \bar{G}_\varepsilon(x)| = \left|\int (\bar{G}_\varepsilon(y) - \bar{G}_\varepsilon(x)) P_\varepsilon(x-y)\,dy\right| \le$$

$$\le |G|_L \int \|x-y\|_2 P_\varepsilon(x-y)\,dy = |G|_L \underbrace{\int \|z\|_2 P_\varepsilon(z)\,dz}_{\mathbb{E}\|Y\|_2 \le \sqrt{\mathbb{E}\|Y\|_2^2} = \sqrt{m\varepsilon}} \le$$

$$\le \sqrt{m\varepsilon}\,|G|_L \xrightarrow{\varepsilon \to 0} 0$$

By Fatou's lemma,

$$\mathbb{E}[e^{\lambda G}] = \mathbb{E}[\lim_{\varepsilon \downarrow 0} e^{\lambda G_\varepsilon}] \overset{\text{Fatou}}{\le} \liminf_{\varepsilon \downarrow 0} \mathbb{E}[e^{\lambda G_\varepsilon}] \le$$

$$\le e^{c\lambda^2 |G_L|^2/2} \cdot \liminf_{\varepsilon \downarrow 0} e^{\lambda \mathbb{E}(G_\varepsilon)}$$

So it is enough to show $\mathbb{E}(G_\varepsilon) \xrightarrow{\varepsilon \downarrow 0} \mathbb{E}G$

Enough that $\mathbb{E}(G_\varepsilon^2)$ is uniformly bounded for ε small (by uniform integrability)

By $(*)$, $\mathbb{E}[[G_\varepsilon - \mathbb{E}(G_\varepsilon)]^2]$ is uniformly bdd. for all ε.

Then enough that $\mathbb{E}(G_\varepsilon)$ is unif bdd. for ε small.

This follows since with Prob. $\frac{3}{4}$ say,

$$|G_\varepsilon - \mathbb{E}(G_\varepsilon)| \le C \;\&\; |G_\varepsilon - G| \le C \;\&\; |G| \le C$$

$\blacksquare$

<u>Back to random matrices:</u>

<u>Reminder:</u>

<u>Hoffman - Wieland lemma:</u>

A, B symmetric (Hermitian) $N \times N$ matrices, $\lambda_1^A \le \dots \le \lambda_N^A$, $\lambda_1^B \le \dots \le \lambda_N^B$

then:
$$\sum_{i=1}^N |\lambda_i^A - \lambda_i^B|^2 \le \mathrm{Tr}[(A-B)^2] = \sum_{i,j=1}^N (a_{ij} - b_{ij})^2$$

<u>Interpretation:</u>

The mapping from the entries (a_{ij}) of a symmetric matrix to the eigenvalues is a Lipshitz with const. ≤ 1.

As a function from $(a_{ij})_{j \ge i}$, it has lipshitz const. $\le \sqrt{2}$

<u>Implication:</u>

If $F: \mathbb{R}^N \to \mathbb{R}$ is Lip. with Lip. const. $|F|_L$, Then the mapping:

$$(a_{ij})_{j \ge i} \longrightarrow F(\lambda_1^A, \dots, \lambda_N^A) \text{ has Lip. const. } \le \sqrt{2}\,|F|_L.$$

(the Y_i & Z_{ij})

In Particular, if the entries of the Wigner matrix X satisfy LSI with Const. C then:

$$P\left(\left|F(\lambda_1^X, \dots, \lambda_N^X) - \mathbb{E}F(\lambda_1^X, \dots, \lambda_N^X)\right| \ge \delta\right) \le 2e^{\frac{-N^2\delta^2}{4C|F|_L^2}}$$

using that if Z satisfies LSI with const. C, then $\frac{1}{\sqrt{n}} Z$ satisfies LSI with Const. $\frac{C}{N}$.