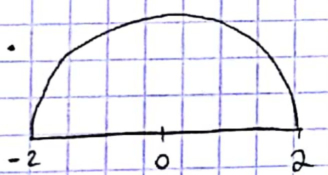
Lecture 3:

7/11/2017

Reminder:

- $X_N = \frac{1}{\sqrt{N}} \begin{pmatrix} Y_1 & Z_{12} & & \\ Z_{12} & Y_2 & & \\ & & \ddots & \\ & & & Y_N \end{pmatrix}$ ← symmetric
- $(Z_{i,j})_{j>i}$ IID } independent
- (Y_j) IID
- $\mathbb{E}(Y_1) = \mathbb{E}(Z_{1,2}) = 0, \quad \mathbb{E}(Z_{1,2}^2) = 1$
- $\forall k, \; r_k = \max\left(\mathbb{E}(|Y_1|^k), \mathbb{E}(|Z_{1,2}|^k)\right) < \infty$
- $\lambda_1^N \leq \dots \leq \lambda_N^N$ eigenvalues
- $L_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i^N}$
- for $f: \mathbb{R} \to \mathbb{R}$: $\langle L_N, f \rangle = \int f(x)\, dL_N(x) = \frac{1}{N} \sum_{i=1}^{N} f(\lambda_i^N)$
- 

  semicircle dist. has density:

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

for each $f \in C_b(\mathbb{R})$: $\forall \varepsilon > 0 \quad P\left(\left|\langle L_N, f\rangle - \langle \sigma, f \rangle\right| > \varepsilon\right) \xrightarrow{N\to\infty} 0$

In proof:

$\bar{L}_N = \mathbb{E}(L_N) \implies \langle \bar{L}_N, f \rangle = \mathbb{E}(\langle L_N, f\rangle)$ for each measurable $f: \mathbb{R} \to [0,\infty)$ & we checked that this indeed defines a measure.

Lemma 1:

$\forall k \geq 0$ integer $\quad \langle \bar{L}_N, x^k \rangle \xrightarrow{N\to\infty} \langle \sigma, x^k \rangle$

Lemma 2:

$\forall k \geq 0, \; \varepsilon > 0 \quad P\left(\left|\langle L_N, x^k\rangle - \langle \bar{L}_N, x^k \rangle\right| > \varepsilon\right) \xrightarrow{N\to\infty} 0$

we saw Lemma 1 & that Lemma 1+2 $\implies$ Thm.

Proof of Lemma 2:

We will show $\mathrm{Var}(\langle Z_N, x^k\rangle) \xrightarrow{N\to\infty} 0$ ...

Recall. $\langle Z_N, x^k\rangle = \sum_{i=1}^{N} (\lambda_i^N)^k = \frac{1}{N}\mathrm{Tr}(X_N^k) =$

$$= \frac{1}{N} \sum_{\substack{i=(i_1,i_2,\dots,i_k) \\ \forall j.\, 1\le i_j\le N}} \underbrace{X_N(i_1,i_2)X_N(i_2,i_3)\dots X_N(i_{k-1},i_k)X_N(i_k,i_1)}_{T_i^N}$$

$$\mathrm{Var}(\langle Z_N, x^k\rangle) = \frac{1}{N^2}\sum_{i,i'} \mathrm{cov}(T_i^N, T_{i'}^N) =$$

$$= \frac{1}{N^2} \sum_{t=1}^{2k} \sum_{\substack{i,i' \\ |\{i_1,\dots i_k, i_1',\dots,i_k'\}| = t}} \mathrm{cov}(T_i^N, T_{i'}^N)$$

union of sets of total size t

Claim:

$\mathrm{cov}(T_i^N, T_{i'}^N) = 0$ when $|\{i_1,\dots,i_k, i_1',\dots,i_k'\}| > k$

Proof:

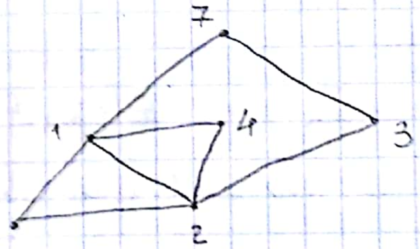
Graph terminology (as in previous lesson):

Associate to i, i' a graph by drawing an edge $\{i_j, i_{j+1}\}$ & $\{i_j', i_{j+1}'\}$ for each $1\le j\le k-1$.

Example:

$i = (2,5,1,4)$

$i' = (3,7,1,2)$



For each edge e, N_e = number of times e is crossed in any direction.

We will use $\mathrm{Cov}(T_i^N, T_{i'}^N) = \mathbb{E}(T_i^N T_{i'}^N) - \mathbb{E}(T_i^N)\mathbb{E}(T_{i'}^N)$.

- If $N_e = 1$ for some edge, then the cov is zero, so we need all $N_e \geq 2$ (הסבר: נובע מות הצלע הזו החוצה מהתחלת והתוחלת שלה מתאפסת ולכן הכל מתאפס)

- If the set of edges of i & i' have no edge in common then $T_i^N, T_{i'}^N$ are independent $\Longrightarrow \mathrm{Cov}(T_i^N, T_{i'}^N) = 0$

Scanned by CamScanner

We will show that with these two observations we can prove the Lemma.

There are k steps in the walk of i & k steps in the walk of i'.

If $\operatorname{cov}(T_i^N, T_{i'}^N) \neq 0$ then $N_e \geq 2$ for all e, so have at most k edges in graph.

הגרף קשיר כי גרף i קשיר וכך גם גרף i', ויש להם קשת משותפת כי אחרת ה- cov היה מתאפס. הגרף קשיר ויש לכל היותר k קשתות ולכן יש לו גם לכל היותר $k+1$ קודקודים.

כמו כן סימנו כי כבר הוכחנו שאם יש יותר מ- $k+1$ הקודקודים בקבוצה $\{i_1,\dots,i_k,i'_1,\dots,i'_k\}$ אז ה- cov מתאפס, מספיק לבדוק עבור $k+1$.

If $\#\text{vertices} = k+1$ then necessarily there are exactly k edges & the graph is a tree. Since the walk of i starts & ends in the same vertex & defines a tree, each edge of it is traversed an even number of times. Thus, the common edge is traversed at least 4 times. Thus, there are at most $k-1$ edges in the graph, so $\#\text{vertices} \leq k$. A contradiction.

With the claim
$$\operatorname{Var}(\langle Z_k, x^k \rangle) = \frac{1}{N^2} \sum_{t=1}^{k} \sum_{\substack{i,i' \\ |\{i_1,\dots,i_k,i'_1,\dots,i'_k\}|=t}} \operatorname{cov}(T_i^N, T_{i'}^N)$$

Now
$$\operatorname{cov}(T_i^N, T_{i'}^N) = \mathbb{E}(T_i^N \cdot T_{i'}^N) - \mathbb{E}(T_i^N)\,\mathbb{E}(T_{i'}^N) =$$
$$= \frac{1}{N^k}\Big[\prod_{e\in E^c} \mathbb{E}(Z_{1,2}^{N_e}) \prod_{e\in E^s} \mathbb{E}(Y_1^{N_e}) - \prod_{e\in E^c} \mathbb{E}(Z_{1,2}^{N_{e,i}})\,\mathbb{E}(Z_{1,2}^{N_{e,i'}}) \cdot$$
$$\cdot \prod_{e\in E^s} \mathbb{E}(Y_1^{N_{e,i}})\,\mathbb{E}(Y_1^{N_{e,i'}})\Big]$$

So $|\operatorname{cov}(T_i^N, T_{i'}^N)| \leq \frac{1}{N^k} G_k$

כי הסכום המכפלים תלוי רק ב- k ולא ב- N. לכן ניתן לכתוב אותו כקבוע G_k.

Hence:

$$Var(\langle \mathcal{L}_N, x^k \rangle) \leq \sum_{t=1}^{k} \frac{C_k}{N^{k+2}} \Big| \underbrace{\{(i,i') \,\big|\, |\{i_1,\dots,i_k,i'_1,\dots,i'_k\}| = t\}}_{} \Big| \leq$$

נחשב כמה זוגות (i,i') כאלו יש. המספר הזה שווה ל-

$N^t t^{2k} \leq N^t k^{2k}$ (כי $t \leq k$). ולכן:

לבחור לְמה שווה כל אחד

לבחור את ה-t שנמצאים בפנים

$$\leq \sum_{t=1}^{k} \frac{C_k k^{2k}}{N^{k+2-t}} \xrightarrow{N\to\infty} 0.$$

$\boxtimes$

Complex Wigner matrices:

- $X_N = \frac{1}{\sqrt{N}} \begin{pmatrix} Y_1 & Z_{12} & \cdots \\ \overline{Z}_{12} & \ddots & \\ \vdots & & Y_n \end{pmatrix} \leftarrow$ Hermition (הרמיטית)
- $(Z_{ij})_{j>i}$ IID complex valued random variables
- (Y_i) IID real valued

independent

- $\mathbb{E}(Y_1) = \mathbb{E}(Z_{1,2}) = \mathbb{E}(Z_{1,2}^2) = 0$
- $\mathbb{E}(|Z_{1,2}|^2) = 1$
- $r_k = \max\left(\mathbb{E}(|Y_1|^k), \mathbb{E}(|Z_{1,2}|^k)\right) < \infty$
- real eigenvalues $\lambda_1^N \leq \dots \leq \lambda_N^N$
- $\mathcal{L}_N = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i^N}$

Wigner's Thm. still holds, & is proved in the same way:

- Lemma 1: $\forall k$, $\langle \mathcal{L}_N, x^k \rangle \xrightarrow{N\to\infty} \langle \sigma, x^k \rangle$
- Lemma 2: $\forall k$ $Var(\langle \mathcal{L}_N, x^k \rangle) \xrightarrow{N\to\infty} 0$

Same proof goes through with the following difference.

$$\langle \mathcal{L}_N, x^k \rangle = \frac{1}{N}\sum_i \mathbb{E}(T_i^N)$$

בהוכחה הקודמת, כאשר ספרנו על קשת פעמיים באותו כיוון. זה היה שקול לספור צלע בכיוונים שונים. עכשיו זה לא שקול כי $X_{ij} \neq X_{ji}$.

$(X_{ij} = \overline{X_{ji}})$

For $\mathbb{E}(T_i^N)$ to be non zero, we need that each edge is crossed at least once in every direction. One complication is that $\mathbb{E}(T_i^N)$ is not exactly the same in every equivalence.

Section 2.1.5 in the book.

Reducing the moment restriction:

Lemma: (Hoffman - Wielandt) (טענה מאוד הע"ע משתנים ע"י שינוי קטן במטריצה)

Let A, B be $N \times N$ symmetric matrices with eigenvalues $\lambda_1^A \le \lambda_2^A \le \dots \le \lambda_N^A$, $\lambda_1^B \le \lambda_2^B \le \dots \le \lambda_N^B$ respectively.

Then: $$\sum_{i=1}^{N} |\lambda_i^A - \lambda_i^B|^2 \le Tr((A-B)^2)$$

(אי השוויון הנ"ל הדוק - נבחר $B = A + \varepsilon I$ ונקבל שוויון)

Proof:

$$Tr((A-B)^2) = Tr(A^2 - AB - BA + B^2) \overset{Tr(AB)=Tr(BA)}{=} \sum_{i=1}^{N}\left[(\lambda_i^A)^2 + (\lambda_i^B)^2\right] - 2\,Tr(AB)$$

on the other hand:

$$\sum_{i=1}^{N} (\lambda_i^A - \lambda_i^B)^2 = \sum_{i=1}^{N}\left[(\lambda_i^A)^2 + (\lambda_i^B)^2\right] - 2\sum_{i=1}^{N} \lambda_i^A \lambda_i^B$$

So we must show that $Tr(AB) \le \sum_{i=1}^{N} \lambda_i^A \lambda_i^B$.

Now $A = U_A D_A U_A^T$ with U_A orthogonal & D_A diagonal.

Similarly, $B = U_B D_B U_B^T$.

So $$Tr(AB) = Tr(U_A D_A \underbrace{U_A^T U_B}_{U = \text{orthogonal}} D_B U_B^T) = Tr(D_A U D_B U^T)$$

on the other hand, $\sum_{i=1}^{N} \lambda_i^A \lambda_i^B = Tr(D_A D_B)$

נרצה למקסם את ה- U ולהראות שהערך המקסימלי כאשר $U = I$

$$Tr(D_A U D_B U^T) = \sum_{i,j} \lambda_i^A \lambda_i^B U_{ij}^2$$

because: $$Tr(XYZW) = \sum_{i_1,i_2,i_3,i_4} X_{i_1,i_2} Y_{i_2 i_3} Z_{i_3,i_4} W_{i_4,i_1}$$

U is orthogonal & therefore: $\sum_{j=1}^{N} U_{ij}^2 = 1$, $\sum_{i=1}^{N} U_{ij}^2 = 1$

Thus $$Tr(AB) \le \max_{\substack{\text{matrices } V \\ V_{ij} \ge 0 \\ \sum_{i=1}^{N} V_{ij} = \sum_{j=1}^{N} V_{ji} = 1}} \left(\sum_{i,j} \lambda_i^A \lambda_i^B V_{ij}\right)$$

$\leftarrow$ Double Stochastic matrices

Claim:

This maximum $= \sum_i \lambda_i^A \lambda_i^B$

Abstract Proof:

We maximize a linear function over a convex space, so the maximum is obtained at the edges. We get that the matrix is a permutation matrix, & the "best" permutation matrix is I.

Proof:

Fix a double stochastic V. Suppose $V_{11} < 1$. We will define a new doubly stochastic $\tilde{V}$ with $\tilde{V}_{ii} \geq V_{ii}$ for all i & having more zero entries in $\tilde{V}_{ij}$ than V_{ij} &

$$\sum_{ij} \lambda_i^A \lambda_j^B \tilde{V}_{ij} \geq \sum_{ij} \lambda_i^A \lambda_j^B V_{ij}$$

(נראה שיהיו יותר אפסים כדי
שנדע שהתהליך מסתיים)

Iterating this proves the claim.

We Define:

$$\begin{array}{c} \\ j - \\ e - \end{array} \begin{array}{c} \begin{array}{cc} \quad m \quad & \quad j \quad \end{array} \\ \begin{pmatrix} & \\ V_{jm} > 0 & V_{jj} < 1 \\ V_{em} & V_{ej} > 0 \end{pmatrix} \end{array}$$

$$\tilde{V}_{jj} = V_{jj} + \varepsilon$$
$$\tilde{V}_{em} = V_{em} + \varepsilon$$
$$\tilde{V}_{jm} = V_{jm} - \varepsilon$$
$$\tilde{V}_{ej} = V_{ej} - \varepsilon$$

נבדוק כי האופרציה שהגדרנו אכן מקיימת את התכונות שדרשנו.

$$\sum_{i,j} \lambda_i^A \lambda_j^B (\tilde{V}_{ij} - V_{ij}) = \varepsilon \left(\lambda_j^A \lambda_j^B + \lambda_e^A \lambda_m^B - \lambda_j^A \lambda_m^B - \lambda_e^A \lambda_i^B \right) =$$

$$= \varepsilon (\lambda_j^A - \lambda_e^A)(\lambda_j^B - \lambda_m^B) \underset{\text{when } j=1}{\geq} 0$$

we will always choose the smallest j s.t. $V_{jj} < 1$.

Thm:

$\forall f \in C_b(\mathbb{R})$. $P\left(|\langle L_N, f\rangle - \langle \sigma, f\rangle| > \varepsilon\right) \xrightarrow[N\to\infty]{} 0$ even when we only assume $r_2 < \infty$ (without asking that $r_k < \infty$ for $k \geq 3$).

Proof:

We will show it for functions f that satisfy:

$$\forall x\ |f(x)| \leq 1, \quad \forall x,y\ |f(x)-f(y)| < |x-y|$$

(Proving for general $f \in C_B(\mathbb{R})$ is an exercise.)

Fix such an f. For given $C > 0$ define:

$$\hat{X}_N(i,j) = X_N(i,j)\, \mathbb{1}_{|\sqrt{N} X_N(i,j)| \leq C} - \mathbb{E}\left(X_N(i,j)\, \mathbb{1}_{|\sqrt{N} X_N(i,j)| \leq C}\right)$$

Now $|\langle L_N, f\rangle - \langle \underbrace{\hat{L}_N}_{\text{empirical eigenvalue measure for } \hat{X}_N}, f\rangle| \leq \frac{1}{N}\sum_{i=1}^{N} |f(\lambda_i^N) - f(\hat{\lambda}_i^N)| \underset{\text{ליפשיצית}}{\leq}$

$$\leq \frac{1}{N}\sum_{i=1}^{N} |\lambda_i^N - \hat{\lambda}_i^N| \underset{\text{קושי שוורץ}}{\leq} \frac{1}{N}\sqrt{N}\sqrt{\sum_{i=1}^{N} |\lambda_i^N - \hat{\lambda}_i^N|^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}|\lambda_i^N - \hat{\lambda}_i^N|^2} \underset{\text{H-W Lemma}}{\leq}$$

$$\leq \sqrt{\frac{1}{N}\mathrm{Tr}\left((X_N - \hat{X}_N)^2\right)}$$

נשים לב כי $\mathrm{Tr}(AA^T) = \sum_{i,j} A_{i,j}^2$ ובפרט

ש- $X_N - \hat{X}_N$ סימטרית נקבל כי:

$$\frac{1}{N}\mathrm{Tr}\left((X_N - \hat{X}_N)^2\right) = \frac{1}{N}\sum_{i,j}\left(X_N(i,j) - \hat{X}_N(i,j)\right)^2 =$$

$$= \frac{1}{N}\sum_{i,j}\left[X_N(i,j)\,\mathbb{1}_{\sqrt{N}|X_N(i,j)| > C} - \mathbb{E}\left(X_N(i,j)\,\mathbb{1}_{\sqrt{N}|X_N(i,j)| > C}\right)\right]^2 = W_N$$

To show that $P(W_N > \varepsilon)$ is small, bound $\mathbb{E}(W_N)$ & use markov.

$$\mathbb{E}(W_N) \underset{(a-b)^2 \leq 2(a^2+b^2)}{\leq} \frac{4}{N^2}\sum_{i,j}\mathbb{E}\left[Z_{i,j}^2\, \mathbb{1}_{|Z_{i,j}| > C}\right] + \left|\mathbb{E}\left(Z_{i,j}\,\mathbb{1}_{|Z_{i,j}| > C}\right)\right|^2$$

The expectations do not depend on N & go to zero as $C \to \infty$ by the dominant Convergance Thm.

All together $\sup_N \mathbb{E}(W_N) \xrightarrow{C \to \infty} 0$

So for large C, $P(|\langle L_N, f\rangle - \langle \hat{L}_N, f\rangle| > \varepsilon) < \varepsilon$.

So, adding a normalization to the entries of $\hat{X}_N$ to make them have Var 1, allows us to use Wiegner's Thm. for $\hat{X}_N$ & conclude that $|\langle \hat{L}_N, f\rangle - \langle \sigma, f\rangle|$ is small as $N \to \infty$.