EE378B Inference, Estimation, and Information Processing

The semicircle law

Andrea Montanari *Lecture 13-14 - Due on 3/8/2021*

Homework should be submitted via Gradescope, by Monday afternoon (unless Monday is a holiday): the code will be communicated by an announcement on Canvas.

For getting credit for the class, you are required to present solutions of some of these homeworks during the first 15 minutes of class starting on 1/20. Please, sign up for (at least) one slot, and be sure that your explanation lasts 15 minutes (or less). For these presentations, you are free to choose whatever format you prefer (slides, typed notes, handwriting, ...).

This week, the presentations will be:

- Monday: Questions (a), (b)
- Wednesday: Questions (c), (d).

Problem

The objective of this homework is to fill-in some of the technical steps in the proof of Wigner's semicircle law that we skipped in class, and to perform some numerical experiments.

Throughout $(W_{ij})_{1\le i<j}$ is an array of i.i.d. random variables with $\mathbb{E}\{W_{ij}\}=0$, $\mathbb{E}\{W_{ij}^2\}=1$, and we denote by $\boldsymbol{W}_n\in\mathbb{R}^{n\times n}$ the simmetric matrix with entries $(W_{ij})_{1\le i<j\le n}$ above the diagonal and $W_{ii}=0$.

(a) *Truncation argument.* Recall that for $K\ge 0$, $W_{ij}^K := W_{ij}\mathbf{1}_{|W_{ij}|\le K}$, and denote by $\boldsymbol{W}_n^K$ the corresponding random matrix. Use Wielandt-Hoffmann inequality and the strong law of large numbers to prove that, almost surely,

$$\limsup_{n\to\infty}\frac{1}{n}\sum_{i=1}^n\Big|\lambda_i\big(\boldsymbol{W}_n/\sqrt{n}\big)-\lambda_i\big(\boldsymbol{W}_n^K/\sqrt{n}\big)\Big|^2\le\mathbb{E}\big\{W_{12}^2\mathbf{1}_{|W_{12}|>K}\big\}=:\Delta_K\,. \tag{1}$$

(b) Let $F_n(x):=\sum_{i=1}^n\mathbf{1}_{x\ge\lambda_i(\boldsymbol{W}_n/\sqrt{n})}/n$, be the distribution of the eigenvaluesof $\boldsymbol{W}_n$, and F_n^K be the analogous distribution for $\boldsymbol{W}_n^K$. Assume that $F_n^K(x)\to F^K(x)$ as $n\to\infty$ for any x. Deduce that, for any $K<\infty$, $\epsilon>0$

$$F^K(x-\epsilon)-\frac{\Delta_K}{\epsilon^2}\le\liminf_{n\to\infty}F_n(x)\le\limsup_{n\to\infty}F_n(x)\le F^K(x+\epsilon)+\frac{\Delta_K}{\epsilon^2}\,. \tag{2}$$

Deduce that, if $\lim_{K\to\infty}F^K(z)=F(z)$ for every $z\in\mathbb{R}$, where F is continuous at x, then

$$\lim_{n\to\infty}F_n(x)=F(x)\,. \tag{3}$$

(c) Prove that the above conditions hold with probability one for F_n^K.

(d) Consider the following symmetric matrix completion problem. Genrate uniformly random orthogonal matrices $\boldsymbol{U}\in\mathbb{R}^{n\times k}$, $n\ge k$, $\boldsymbol{U}^{\mathsf{T}}\boldsymbol{U}=\boldsymbol{I}_{k\times k}$. (By the way, how can you generate such random matrices). Let

$\boldsymbol{X} = \boldsymbol{U}\boldsymbol{U}^{\mathsf{T}}$. Further, let $\boldsymbol{Y} \in \mathbb{R}^{n\times n}$ be a symmetric matrix with $Y_{ii} = 0$ and

$$Y_{ij} = \begin{cases} X_{ij} & \text{with probability } p_n, \\ 0 & \text{with probability } 1 - p_n, \end{cases} \tag{4}$$

for $i < j$. For $i > j$, we set $Y_{ij} = Y_{ji}$.

Fix $k = 4$ and setting $p_n = (2\log n)/n$, generalte random matrices $\boldsymbol{Y}$ distributed as above for $n \in \{500, 1000, 2000, 4000\}$.

- For each n, compute the eigenvalues of $\boldsymbol{Y}$ and plot their histogram. Compare it with what would be expected on the basis of Wigner's semicircle's law.
- Discuss the differences between the semicircle's law and what you observe.
- Can you apply Wigner's law as proved in class to this example? Justify your answer.