Concentration of the empirical spectral distribution of random matrices with dependent entries

Bartłomiej Polaczyk

Institute of Mathematics, University of Warsaw

Banacha 2, 02-097 Warsaw, Poland

b.polaczyk@mimuw.edu.pl

Abstract

We investigate concentration properties of spectral measures of Hermitian random matrices with partially dependent entries. More precisely, let X_n be a Hermitian random matrix of the size $n \times n$ that can be split into independent blocks of the size at most $d_n = o(n^2)$. We prove that under some mild conditions on the distribution of the entries of X_n, the empirical spectral measure $L_n^{\frac{1}{\sqrt{n}}X_n}$ of $\frac{1}{\sqrt{n}}X_n$ concentrates around its mean, i.e. $\rho(L_n^{\frac{1}{\sqrt{n}}X_n}, \mathbb{E}L_n^{\frac{1}{\sqrt{n}}X_n}) \to_{\mathbb{P}} 0$, where ρ is any metric that metrizes weak convergence of probability measures.

1. Introduction

Let X_n be a random $n \times n$ Hermitian matrix. All eigenvalues $\lambda_1 \le \ldots \le \lambda_n$ of X_n lie on the real line and thus we may consider its **empirical spectral distribution** (ESD) being a probability measure on $\mathbb{R}$ given by the formula

$$L_n^{X_n} := \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i}.$$

Since X_n is random, then so is $L_n^{X_n}$. One can thus consider its expected value – a deterministic probability measure $\mathbb{E}L_n^{X_n}$ s.t. for every compactly supported continuous function f

$$\int f \, d\mathbb{E}L_n^{X_n} = \mathbb{E}\int f \, dL_n^{X_n}$$

(the existence of $\mathbb{E}L_n$ follows from Riesz-Markov-Kakutani representation theorem).

When dealing with sequences of **random probability measures** we distinguish (at least) three distinct modes of convergence. Let ρ be any metric that metrizes weak convergence of probability measures (e.g. Prokhorov's or Lévy's metric). We say that μ_n converges to μ (weakly)

$$\begin{aligned} \text{almost surely, } & \mu_n \Rightarrow \mu, \text{ if } \rho(\mu_n, \mu) \to 0 \;\; a.s., \\ \text{in probability, } & \mu_n \overset{\mathbb{P}}{\Rightarrow} \mu, \text{ if } \rho(\mu_n, \mu) \overset{\mathbb{P}}{\to} 0, \\ \text{in expectation, } & \mathbb{E}\mu_n \Rightarrow \mu, \text{ if } \rho(\mathbb{E}\mu_n, \mu) \to 0. \end{aligned}$$

The alternative characterization is often more handy:

$$\begin{aligned} \mu_n \Rightarrow \mu &\iff \forall_{f \in C_{c,L}} \int f \, d\mu_n \to \int f \, d\mu \;\; a.s., \\ \mu_n \overset{\mathbb{P}}{\Rightarrow} \mu &\iff \forall_{f \in C_{c,L}} \int f \, d\mu_n \overset{\mathbb{P}}{\to} \int f \, d\mu, \\ \mathbb{E}\mu_n \Rightarrow \mu &\iff \forall_{f \in C_{c,L}} \mathbb{E}\int f \, d\mu_n \to \int f \, d\mu, \end{aligned}$$

where $C_{c,L}$ is the set of all continuous compactly supported 1-Lipschitz real functions. It follows that

$$\left[\mu_n \Rightarrow \mu\right] \implies \left[\mu_n \overset{\mathbb{P}}{\Rightarrow} \mu\right] \implies \left[\mathbb{E}\mu_n \Rightarrow \mu\right].$$

A milestone that began random matrix theory is the following result from [Wig]. It says that whenever the entries of X_n are i.i.d. (up to the symmetry constrain) with mean 0 and variance 1, then

$$L_n^{\frac{1}{\sqrt{n}}X_n} \Rightarrow \sigma, \;\; \text{with } \; \sigma(x) = \frac{1}{2\pi}\sqrt{4 - x^2}\,\mathbf{1}_{|x|<2}.$$

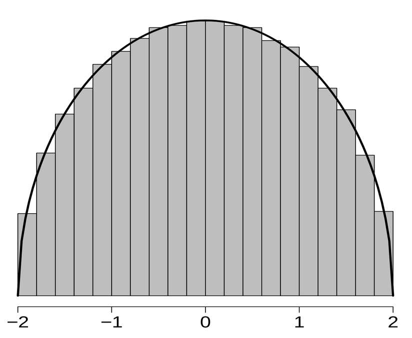

Figure 1: Distribution of the eigenvalues of 2000×2000 normalized random matrix with i.i.d. entries together with σ (black line).

Much is known about the limiting behavior of the ESD of random matrices with **independent** entries but in many applications one has to deal with models of matrices having **dependent** entries. In these cases the analysis is usually much more involved and results – weaker (see e.g. recent work [SchSch], where authors analyze Anderson lattice model and obtain convergence **in expectation** to σ).

2. Main Result

Our main theorem can be stated as follows.

Theorem 2.1. *Assume that X_n is a sequence of random Hermitian matrices s.t. for each n the matrix X_n can be split into stochastically independent blocks of size at most d_n, where $d_n = o(n^2)$ (that is the size of each block is asymptotically dominated by the total number of entries).*

$$\begin{bmatrix} X_{11} & X_{12} & X_{13} & X_{14} \\ X_{21} & X_{22} & X_{23} & X_{24} \\ X_{31} & X_{32} & X_{33} & X_{34} \\ X_{41} & X_{42} & X_{43} & X_{44} \end{bmatrix}$$

Figure 2: Each color corresponds to each block of a matrix. Distinct blocks are stochastically independent random vectors of length $\le d_4$.

If the family $\{|(X_n)_{ij}|^2\}_{1 \le i,j \le n \in \mathbb{N}}$ is uniformly integrable, then for any metric ρ that metrizes weak convergence of probability measures

$$\rho(L_n^{\frac{1}{\sqrt{n}}X_n}, \mathbb{E}L_n^{\frac{1}{\sqrt{n}}X_n}) \overset{\mathbb{P}}{\to} 0.$$

In particular, if $\mathbb{E}L_n^{\frac{1}{\sqrt{n}}X_n} \Rightarrow \mu$, then $L_n^{\frac{1}{\sqrt{n}}X_n} \overset{\mathbb{P}}{\Rightarrow} \mu$.

In general we cannot get better than $d_n = o(n^2)$. To see that, consider two random matrix ensembles X_n, Y_n, whose ESDs converge a.s. to distinct limits μ and ν. Set

$$Z_n = \epsilon X_n + (1 - \epsilon) Y_n,$$

where $\mathbb{P}(\epsilon = 0) = \mathbb{P}(\epsilon = 1) = 0.5$ and ϵ is independent of all X_n and Y_n. Then $d_n = n^2$, $L_n^{Z_n} \Rightarrow (\epsilon\mu + (1-\epsilon)\nu)$ a.s. and thus $\rho(L_n^{Z_n}, \mathbb{E}L_n^{Z_n})$ cannot converge in probability to zero for any metric ρ that metrizes weak convergence of probability measures. The above example can be easily modified to get $d_n = \alpha n^2$ for any $\alpha \in (0, 1)$.

3. Tools used

The proof of Theorem 2.1 is based on the argument from [GuZe]. The main tools are the following theorem due to Talagrand and a standard lemma from random matrix theory.

Theorem 3.1 ([Tal])**.** *Let $\{Y_k \in \mathbb{R}^{n_k}\}_{k=1}^r$ be stochastically independent random vectors s.t. $\||Y_k|\|_\infty < D$ for every k. Then there exists a universal constant c s.t. for any convex 1-Lipschitz function $F : \mathbb{R}^{n_1 + \ldots + n_r} \to \mathbb{R}$*

$$\mathbb{P}\left(|F(Y_1, \ldots, Y_r) - \mathbb{E}F(Y_1, \ldots, Y_r)| > t\right) \le 4\exp(-ct^2/D^2).$$

Lemma 3.2. *For any real convex 1-Lipschitz function h, the mapping*

$$X_n \overset{F}{\mapsto} \int h \, dL_n^{\frac{1}{\sqrt{n}}X_n}$$

is convex and $\frac{1}{n}$-Lipschitz w.r.t. Hilbert-Schmidt norm of a matrix.

4. Sketch of the proof

0. We will provide a proof of the theorem in the simplified case, where the entries of X_n are uniformly bounded and the ESD of $\frac{1}{\sqrt{n}}X_n$ converges, i.e. we assume that $\||(X_n)_{ij}|\|_\infty \le K$ and $\mathbb{E}L_n^{\frac{1}{\sqrt{n}}X_n} \Rightarrow \mu$. The first inequality implies that X_n can be decomposed into blocks $Y_1, \ldots, Y_r$, s.t. $\||Y_k|\|_\infty \le K\sqrt{d_n}$.
1. Fix any $f \in C_{c,L}$ and $t > 0$. Our aim is to show that
$$(\clubsuit) := \mathbb{P}\left(\left|\int f \, dL_n^{\frac{1}{\sqrt{n}}X_n} - \mathbb{E}\int f \, dL_n^{\frac{1}{\sqrt{n}}X_n}\right| > t\right) \to 0.$$
It will imply that $\int f \, dL_n^{\frac{1}{\sqrt{n}}X_n} \overset{\mathbb{P}}{\to} \int f \, d\mu$ and the result will follow by the triangle inequality.
2. Assume that the support of f is contained in the interval $[-R, R]$. We approximate f with f_Δ – a combination of at most $\kappa := 4\lceil R/\Delta \rceil$ **convex**, 1-Lipschitz functions (h_l) s.t. $\|f - f_\Delta\|_\infty < \Delta$. The construction is recursive with $f_\Delta \equiv 0$ for $x \le -R$ and
$$f_\Delta(x) = \sum_{s=0}^{2\lceil R/\Delta \rceil} \underbrace{(2\mathbf{1}_{\{f(-R+(s+1)\Delta) > f_\Delta(-R+s\Delta)\}} - 1)g(x - s\Delta)}_{:=g_s(x)},$$
where $g(x) = \max(0, x) - \max(0, x - \Delta)$. Clearly g_s is a difference of two **convex** functions – all of which define the family $\{h_l\}_{l=1}^{\kappa}$.

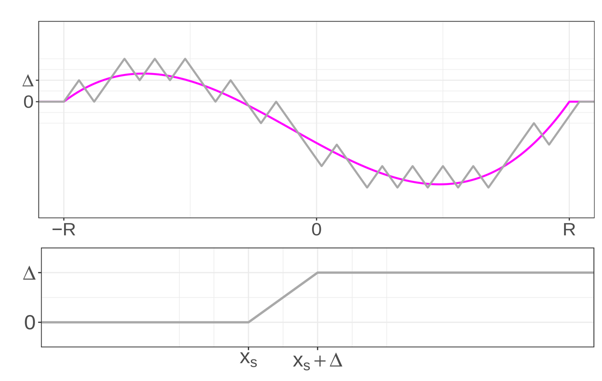

Figure 3: We define f_Δ (upper figure, grey) recursively: in s-th step adding or subtracting g_s (lower figure).

3. Set $\Delta \le \frac{t}{3}$, then
$$\begin{aligned} (\clubsuit) &\le \mathbb{P}\left(\left|\int f_\Delta \, dL_n^{\frac{1}{\sqrt{n}}X_n} - \mathbb{E}\int f_\Delta \, dL_n^{\frac{1}{\sqrt{n}}X_n}\right| > \frac{t}{3}\right) \\ &\le \kappa \sup_{1 \le l \le \kappa} \mathbb{P}\left(\left|\int h_l \, dL_n^{\frac{1}{\sqrt{n}}X_n} - \mathbb{E}\int h_l \, dL_n^{\frac{1}{\sqrt{n}}X_n}\right| > \frac{t}{3\kappa}\right) \\ &\le 4\kappa \exp\left(-\frac{ct^2}{K^2 \cdot 9\kappa^2} \cdot \frac{n^2}{d_n}\right), \end{aligned}$$
where the last inequality follows from Theorem 3.1 and Lemma 3.2 with $F_s(X) = \int h_s \, dL_n^{\frac{1}{\sqrt{n}}X_n}$. □

5. Acknowledgment

The results presented in this poster are a part of a master's thesis written at the University of Warsaw under supervision of prof. Radosław Adamczak.

References

[GuZe] A. Guionnet and O. Zeitouni. Concentration of the spectral measure for large matrices. *Electron. Comm. Probab.*, 5:119–136 (electronic), 2000.

[SchSch] J. Schenker, H. Schulz-Baldes. Semicircle law and freeness for random matrices with symmetries or correlations. *Math. Res. Lett.*, 12:10.4310, 2005.

[Wig] E. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. *Ann. Math.*, 62:548–564, 1955.

[Tal] M. Talagrand. Concentration of measure and isoperimetric inequalities in product spaces. *Inst. Hautes Études Sci. Publ. Math.*, 81:73–205, 1995.