Lévy-Khintchine random matrices

Paul Jung
University of Alabama Birmingham

September 21, 2014
University of Cincinnati

Motivation

- Wigner matrices ('55, '58).
- Heavy tailed matrices have i.i.d. entries (up to symmetry) with infinite variance. Cizeau, Bouchaud, Soshnikov, Ben Arous, Guionnet (08); Bordenave, Caputo, Chafai ('11).
- Adjacency matrices of Erdös-Rényi graphs with $p = 1/n$. Rogers, Bray, Zakharevich ('06), Bordenave and Lelarge ('10).
- General symmetric matrices with symmetric i.i.d. entries: Sum of a row converges weakly as $n \to \infty$. Limits are infinitely divisible $ID(\sigma^2, d, \nu)$.

Annals of Mathematics 1958)

ANNALS OF MATHEMATICS
Vol. 67, No. 2, March, 1958
Printed in Japan

ON THE DISTRIBUTION OF THE ROOTS OF CERTAIN SYMMETRIC MATRICES

BY EUGENE P. WIGNER

(Received September 19, 1957)

The present article is concerned with the distribution of the latent roots (characteristic values) of certain sets of real symmetric matrices of very high dimensionality. Its purpose is to point out that the distribution law obtained before[1] for a very special set of matrices is valid for much more general sets. The dimension of the matrices will be denoted by N, the matrix elements by v_{ij}. These are real. The condition of symmetry is

$$v_{ij} = v_{ji} \tag{1}$$

The matrix elements v_{ij} of the set of matrices are distributed according to the following laws:

(a) The distribution $p_{ij}(v_{ij})$ of the v_{ij} for $i \leqq j$ are independent. In other words, there are no statistical correlations between the matrix elements, except for the condition of symmetry.

(b) The distribution law for each v_{ij} is symmetric.

(c) The distribution laws for all v_{ij} are such that all moments of v_{ij} exist and have an upper bound which is independent of i and j. Because of (b) the odd moments all vanish.

(d) The second moment of all v_{ij} is the same and will be denoted by m^2.

Actually, the last condition can be relaxed considerably so that it holds only for the large majority of the matrix elements. However, this point will not be pursued further. The preceding postulates can be summarized by the postulate that the fraction of the matrices of the set for which the i, j matrix element is within unit interval at v_{ij} is

$$P(v_{11}, v_{12}, v_{13}, \cdots, v_{NN}) = \prod_{i \leq j} p_{ij}(v_{ij}) \tag{2}$$

where

$$\int p_{ij}(v)dv = 1 \tag{2a}$$

and

[1] E. P. WIGNER, Ann. of Math., 62 (1955), 548. The title of the relevant section is *Random Sign Symmetric Matrix*, pp. 552-557. The distribution of the roots of "singly bordered" matrices in which only the diagonal elements are subject to random fluctuations was given by F. J. DYSON, Phys. Rev., 92 (1953), 1331. Cf. also H. SCHMIDT, ibid. 105 (1957), 425.

$$p_{ij}(v) = p_{ij}(-v) \tag{2b}$$

$$\int p_{ij}(v)v^n dv \leqq B_n \tag{2c}$$

$$\int p_{ij}(v)v^2 dv = m^2 . \tag{2d}$$

We consider that the distribution functions p_{ij} are defined for all i, j and that the bound B_n is independent of i and j. All integrals are to be extended from $-\infty$ to ∞.

Under the conditions enumerated, crudely speaking, the fraction of roots within unit interval at x becomes

$$\sigma(x) = \frac{(4Nm^2 - x^2)^{1/2}}{2\pi Nm^2} \qquad \text{for } x^2 < 4Nm^2 \tag{3}$$

and

$$\sigma(x) = 0 \qquad \text{for } x^2 > 4Nm^2 \tag{3a}$$

as N grows beyond all limits. The distribution law (3) was stated before only for the case in which all p_{ij} for $i < j$ are equal and gave the probability 1/2 to the values m and $-m$ of v_{ij} for $i \neq j$ and the probability 1 to the value 0 of v_{ii}. Note that condition (d) is not fulfilled in this case except in the sense of the remark after the statement of that condition.

The theorem can be stated more accurately as follows. Denote by $S_{\alpha,\beta}(v, N)$ (where v, N is an abbreviation for all v_{ij} with $i \leqq j \leqq N$) the number of roots of the N dimensional symmetric matrix $\| v_{ij} \|$ which lie between $\alpha\sqrt{N}$ and $\beta\sqrt{N}$. Then, if the distribution P of the v_{ij} satisfies the conditions given, the fraction of the roots between $\alpha\sqrt{N}$ and $\beta\sqrt{N}$

$$N^{-1}\int \cdots \int P(v, N)S_{\alpha,\beta}(v, N)dv_{11} \cdots dv_{NN} \to \int_\alpha^\beta \frac{(4m^2 - \xi^2)^{1/2}d\xi}{2\pi m^2} \tag{4}$$

as $N \to \infty$ if $-2m < \alpha < \beta < 2m$. If $\alpha < \beta < -2m$ or if $2m < \alpha < \beta$, the left side of (4) tends to zero as $N \to \infty$. Note that the theorem gives the distribution of the roots of sequences of sets of matrices, the matrices of successive sets of the sequence being obtained, from the matrices of the preceding set, by augmenting the matrices with further rows and columns. The distribution of the elements in these added columns is subject, apart from the two conditions of symmetry (1) and (2b), only to the conditions (2c) and (2d). This shows that the distribution of roots depends, under the conditions stated, only on the second moment of the matrix elements.

The heuristic proof given for the special case considered before[1] applies equally under the more general conditions here specified.

PRINCETON UNIVERSITY

Normalization for Wigner matrices

- Empirical (normalized) measure of eigenvalues $e_j(\omega) \in \mathbb{R}$:

$$\frac{1}{n}\sum_{j=1}^{n} \delta_{e_j} = \mathrm{ESD}_n.$$

- To *normalize the entries* note that

$$\mathbf{E}(\text{Second Moment}(\mathrm{ESD}_n)) = \mathbf{E}\frac{1}{n}\,\mathrm{Tr}(A_n^2) = \mathbf{E}\frac{1}{n}\sum_{i,j} a_{ij}a_{ji} = n\mathbf{E}a_{ij}^2.$$

- So we need

$$\mathbf{E}a_{ij}^2 \sim \frac{1}{n}.$$

- Instead of normalizing, change the distribution as n varies:

$$a_{ij} = a_{ji} \sim \text{Bernoulli}\,(\lambda/n) \quad \text{so that} \quad \mathbf{E}a_{ij}^2 = \lambda/n.$$

Normalization for Wigner matrices

- Empirical (normalized) measure of eigenvalues $e_j(\omega) \in \mathbb{R}$:

$$\frac{1}{n}\sum_{j=1}^{n} \delta_{e_j} = \mathrm{ESD}_n.$$

- To *normalize the entries* note that

$$\mathbf{E}(\text{Second Moment}(\mathrm{ESD}_n)) = \mathbf{E}\frac{1}{n}\,\mathrm{Tr}(A_n^2) = \mathbf{E}\frac{1}{n}\sum_{i,j} a_{ij}a_{ji} = n\mathbf{E}a_{ij}^2.$$

- So we need

$$\mathbf{E}a_{ij}^2 \sim \frac{1}{n}.$$

- Instead of normalizing, change the distribution as n varies:

$$a_{ij} = a_{ji} \;\sim\; \text{Bernoulli}\,(\lambda/n) \quad \text{so that} \quad \mathbf{E}a_{ij}^2 = \lambda/n.$$

Main results

Suppose each A_n has i.i.d. entries up to self-adjointness satisfying:

$$\lim_{n\to\infty} \sum_{j=1}^{n} A_n(i,j) \stackrel{d}{=} ID(\sigma^2, d, \nu).$$

J. (2014)

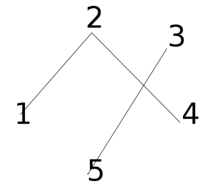
- ESD$_n$ a.s. weakly converge to a symm. prob. meas. μ_∞.
- μ_∞ is the expected spectral measure for vector δ_{root} of a self-adjoint operator on $L^2(G)$.

(Spectral measure for v is defined as $d\langle v, E(t)v\rangle$)

- Wigner matrices: $G = \mathbb{N}$
- Sparse matrices: G is a Poisson Galton-Watson tree

- We need $\mathbf{E} a_{ij}^2 \sim \frac{\lambda}{n}$.
- Adjacency matrices of Erdős-Rényi graphs

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$



Idea of proof

(1) As rooted graphs, Erdős-Rényi(λ/n) *locally converge* to a branching process with a Poiss(λ) offspring distribution.

(2)

Bordenave-Lelarge (2010)

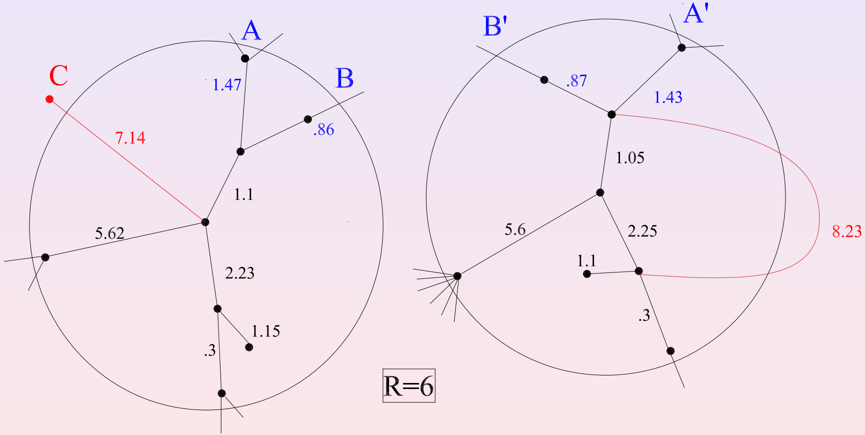
If $G_n[1] \Rightarrow G_\infty[1]$, then one has strong resolvent convergence: for all $z \in \mathbb{C}_+$,

$$(zI - A_n)^{-1}_{11} \to (zI - A_\infty)^{-1}_{11}$$

(3)

$$\mathbf{E}(zI - A_n)^{-1}_{11} = \mathbf{E}\frac{\mathrm{Tr}(zI - A_n)^{-1}}{n} = \int \frac{1}{z-x} d\mathbf{E}(\mathrm{ESD}_n)$$

$\epsilon = 1/6$-close graphs

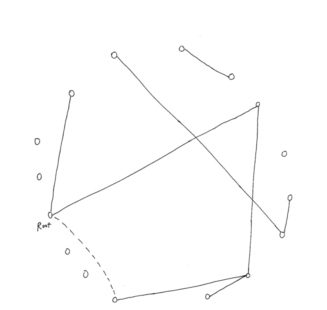


Local weak limits of Erdős-Rényi graphs

- $a_{ij} \sim \text{Bernoulli}(\lambda/n)$ so the number of offspring is Poisson(λ).
- Fix k, an offspring in generation bigger than 1, the probability that it's also a direct offspring (genereation 1) is:

$$\mathbf{P}(1 \sim k) = 1/n \to 0.$$

Local weak convergence to a Poiss(λ) branching process



Weighted-edges case when $\sigma^2 = 0$, $d = 0$: Aldous' PWIT

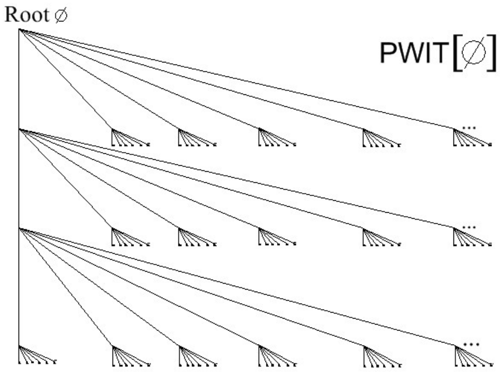


Figure 2: Each [tree symbol] represents a copy of the PWIT. Weights on offspring edges from any vertex are determined by a Poisson process.

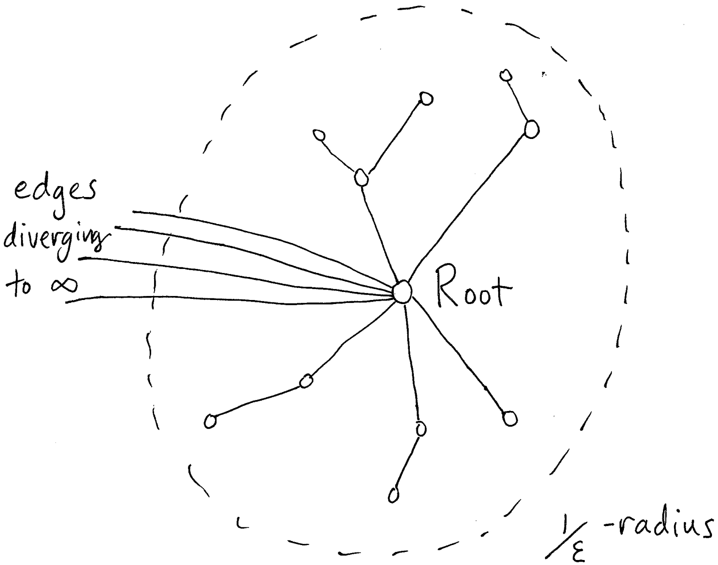
- By Lévy-Itō decomposition, write $A_n = G_n + L_n$
- Local weak convergence implies strong resolvent convergence when $\sigma^2 = 0$ [handles (L_n)].
- Voiculescu's theorem says (G_n) and (L_n) are asymptotically free.
- The LSD of (A_n) is the free convolution of the LSDs of (G_n) and (L_n).

What about σ^2 and d?

- Interlacing handles drift (rank one perturbation).
- For the step in the proof where LWC $\Rightarrow$ Strong resolvent conv. we need

$$\lim_{\varepsilon\searrow 0}\lim_{n\to\infty}\sum_{j=1}^{n}|a_{1j}|^2 1_{\{|a_{1j}|^2\le\varepsilon\}}=0.$$

Problem: edges diverging to infinity



The Poisson weighted infinite skeleton tree

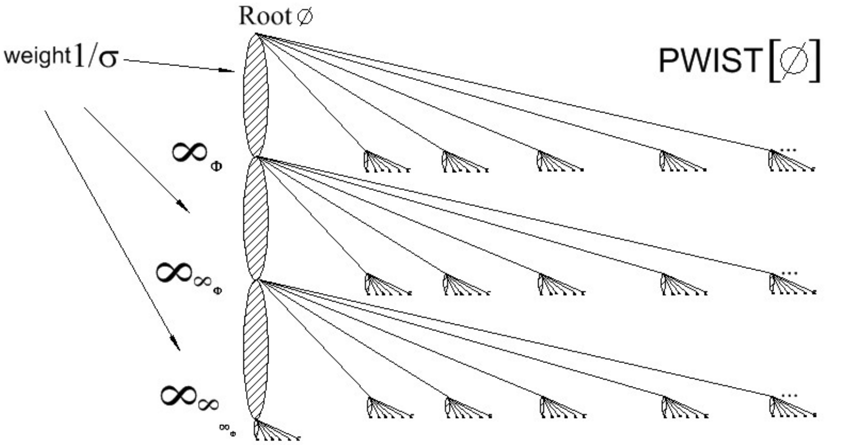


Figure 3: Each [copy symbol] represents a copy of the PWIST. Weights on cords to infinities are deterministic. All other weights are random and determined by Poisson processes as before.

Cords to infinity: $\sigma^2 > 0$

- Distance = resistance on electric networks, and resistance = 1/conductance
- The conductance of each parallel edge is “zero”; however, their collective effective conductance is σ and the effective resistance is $1/\sigma$.
- Identifying all edges with small conductance to one single point we get that

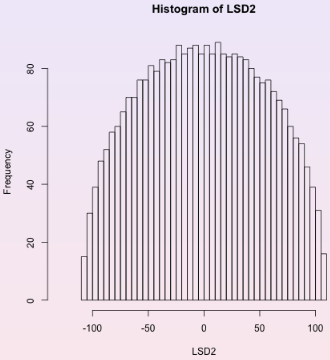
$$\lim_{\varepsilon \searrow 0} \lim_{n\to\infty} \sum_{j=1}^{n} |a_{1j}|^2 1_{\{|a_{1j}|^2 \leq \varepsilon\}} = 0.$$

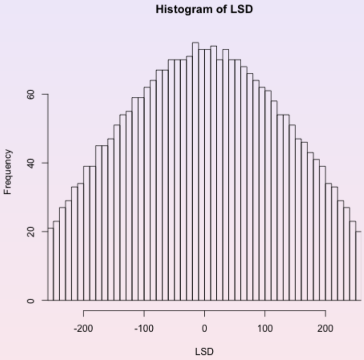
Wigner matrices: vacuum state of the free Fock space

We can handle infinite second moments in the Gaussian domain of attraction.

When there is no Levy measure, the PWIST is the half-line N.
It is well-known that the spectral measure at the root is semi-circle.

Semicircle Pictures





Schur complement formula

Corollary (J. 2014): For $z \in \mathbb{C}_+$, $\quad R_{jj}(z) \overset{d}{=} (A_\infty - zI)^{-1}_{11}$ satisfies

$$R_{00}(z) \overset{d}{=} -\left(z + \sigma^2 R_{11}(z) + \sum_{j \geq 2} a_j^2 R_{jj}(z) \right)^{-1}$$

where $\{a_j\}$ are arrivals of an independent Poisson(ν) process.

Thanks for your attention!

[AS04] David Aldous and J. Michael Steele.
The objective method: Probabilistic combinatorial optimization and local weak convergence.
In *Probability on discrete structures*, pages 1–72. Springer, 2004.

[BAG08] Gérard Ben Arous and Alice Guionnet.
The spectrum of heavy tailed random matrices.
Communications in Mathematical Physics, 278(3):715–751, 2008.

[BCC11a] Charles Bordenave, Pietro Caputo, and Djalil Chafai.
Spectrum of large random reversible Markov chains: heavy-tailed weights on the complete graph.
The Annals of Probability, 39(4):1544–1590, 2011.

[BL10] Charles Bordenave and Marc Lelarge.
Resolvent of large random graphs.
Random Structures & Algorithms, 37(3):332–352, 2010.

[GL09] Adityanand Guntuboyina and Hannes Leeb.
Concentration of the spectral measure of large Wishart matrices with dependent entries.
Electron. Commun. Probab, 14(334-342):4, 2009.

[Zak06] Inna Zakharevich.
A generalization of Wigner's law.
Communications in Mathematical Physics, 268(2):403–414, 2006.