The largest eigenvalue of Hermitian Random Matrices: a moment approach

S. Péché

03/25-30/2007, UC Berkeley,

Workshop on Free Probability and Large N limit

Plan

- I. Introduction
  Motivations for the study of the largest eigenvalue of random matrices.

- II. Standard random matrices.
  A review of known (universality) results.

- III. Soshnikov's approach: Wigner random matrices.

- IV. Extension to other ensembles of random matrices
  A. Non symmetrically distributed entries.
  B. Deformed ensembles.
  C. Sample covariance matrices.

- V. Concluding remarks.

Principal Component Analysis

A portfolio $\mathcal{P}$ of N assets with weight w_i, $i = 1, \ldots, N$.
Σ the covariance matrix of the returns.

$$\text{Daily variance of return to be minimized : } R^2 = \sum_{i,j=1}^{N} w_i w_j \Sigma_{ij}.$$

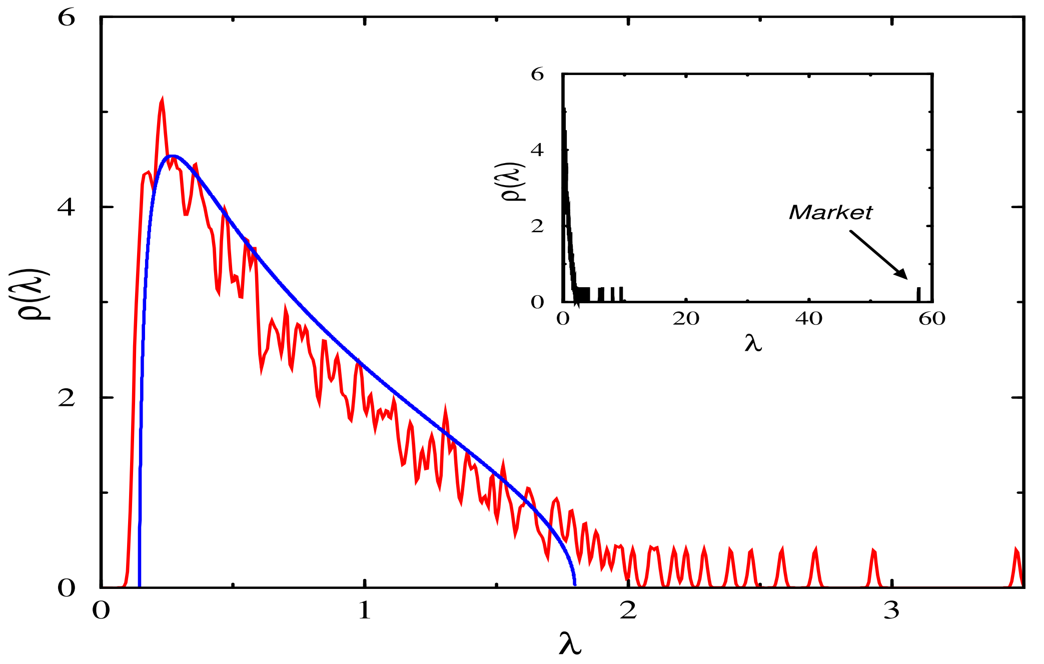
Optimal $\mathcal{P}$: maximal weight on eigenvectors associated to the smallest eigenvalues of Σ.

Problems : Given Y, the $N \times p$ matrix of returns observed during a length p period and the sample covariance matrix YY^*,

1. estimate for Σ unknown in general
2. find a way to erase the effect of large eigenvalues.
3. Principal component analysis in mathematical statistics.

Largest eigenvalue and finance

Bouchaud-Potters-Laloux Financial Applications of RMT (2005)



Empirical eigenvalue density for 406 stocks from the S&P 500, and fit using the MP distribution. Note the presence of one very large eigenvalue.

Random matrices : basic model

Let μ (resp.μ') be a probability distribution on $\mathbb{C}$ (resp. on $\mathbb{R}$) with finite variance.

- a $N \times N$ Hermitian random matrix

$$H_N = \frac{1}{\sqrt{N}}(H_{ij}), \;\; H_{ij}, i \leq j \text{ i.i.d. of distribution } \mu \text{ } (\mu' \text{ on the diagonal})$$

  Archetypical ensemble : GUE $\mu = \mathcal{N}(0,1)$ (complex) and $\mu' = \mathcal{N}(0,1)$ (real).

- a $N \times N$ complex sample covariance matrix :
  $\Sigma > 0$ a $N \times N$ deterministic covariance matrix , $Y = \Sigma^{1/2} X$ (X : $N \times p(N)$, $p(N)/N \to \gamma \in [0, \infty]$),

$$M_N = \frac{YY^*}{N}, \; X = (X_{ij}), \text{ matrix with i.i.d. entries of distribution } \mu.$$

Question: behavior of extreme eigenvalues of such random matrices as $N \to \infty$?

Standard Hermitian random matrices

$H_{ij}, i \leq j$, i.i.d. with distribution μ (or μ' on the diagonal) such that

$$\int x d\mu = 0, \int |x|^2 d\mu = \sigma^2, \int |x|^2 d\mu' < \infty.$$

$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$ the eigenvalues of $H_N = \frac{1}{\sqrt{N}} H$, $\mu_N = \frac{1}{N} \sum_{i=1}^{N} \delta_{\lambda_i}$.

Theorem Wigner (55):

A.s. if μ admits moments of order 4,

$$\lim_{N \to \infty} \mu_N = \sigma_{sc}, \text{ with } \frac{d\sigma_{sc}}{dx} = \frac{1}{2\pi\sigma^2} \sqrt{4\sigma^2 - x^2}\, 1_{[-2\sigma, 2\sigma]}(x).$$

Also holds for real symmetric matrices.

Sample covariance matrices: Marchenko-Pastur distribution.

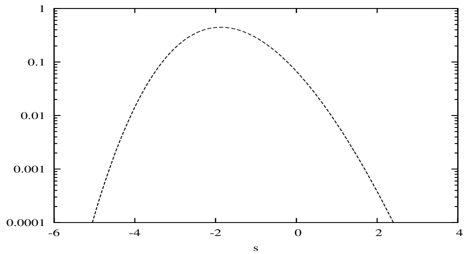
Behavior of extreme eigenvalues?

The largest eigenvalue of Hermitian ensembles

A.s. limit : Bai-Yin (87), Geman (80) Bai-Yin-Krishnaiah (88)
If $\int |x|^4 d\mu(x) < \infty$ and $\int |x|^2 d\mu'(x) < \infty$, then $\lim \lambda_1 = 2\sigma$ a.s.

Fluctuations : Tracy-Widom (94)
λ_1 the largest value of the GUE i.e. $\mu = \mathcal{N}(0, \sigma^2)$, $\mu' = \mathcal{N}(0, \sigma^2)$,

$$\lim_{N\to\infty} P\Big(N^{2/3}\left(\frac{\lambda_1}{2\sigma} - 1\right) \leq x\Big) = F_2^{TW}(x),$$ Tracy Widom distribution.



$F_2^{TW}(x) = \det(I - A_x)$, A_x operator on $L^2(x, \infty)$ with the so-called Airy kernel

$$Ai(u, v) = \frac{Ai(u)Ai'(v) - Ai'(u)Ai(v)}{u - v}.$$

Universality results: the moment approach.

• **Non Gaussian Wigner matrices.** Soshnikov (99):
μ non Gaussian symmetric distribution with sub-Gaussian tails

$$\exists C>0,\ \forall k>0,\ \int |x|^{2k} d\mu(x) \le (Ck)^k \text{ and } \int |x|^2 d\mu(x) = \sigma^2\ (\star).$$

$$\lim_{N\to\infty} P\Big(N^{2/3}\left(\frac{\lambda_1}{2\sigma}-1\right) \le x\Big) = F_2^{TW}(x).$$

• Previous universality results: Invariant Ensembles
Theorem: Deift and al. (99):

$$dP_N(H_N) \propto \exp\{-N\mathrm{Tr}V(H_N)\}dH_N, \text{ where } dH_N \text{Lebesgue measure on } \mathcal{H}_N$$

For a wide class of V, $\exists\, C,\ u_+>0$, $\lim_{N\to\infty} P\left(CN^{2/3}(\lambda_1-u_+) \le x\right) = F_2^{TW}(x)$.

A review of Soshnikov's method

- Gaussian ensembles: $\lambda_1 = 2\sigma + \xi N^{-2/3}$ with $\xi \sim F_2^{TW}$

- If one computes
$$m_k^N(t_1, \ldots, t_k) = \mathbb{E} \prod_{i=1}^{k} Tr \left(\frac{H_N}{2\sigma} \right)^{[t_i N^{2/3}]},$$
for any k, one should find something like the Laplace transform of the joint distribution of largest eigenvalues.

- Instead of computing the asymptotics of m_k^N, show that
$$m_k^N(t_1, \ldots, t_k) = m_k^N(GUE)(t_1, \ldots, t_k)(1 + o(1)).$$

- One can then deduce that the joint distribution of the largest eigenvalues of H_N exhibit Tracy-Widom fluctuations.

Wigner random matrices and Dyck paths

$$\mathbb{E}\left[N^{s_N/2}\mathrm{Tr}H_N^{2s_N}\right] = \sum_{i_o,\dots i_{2s_N-1}} \mathbb{E}H_{i_oi_1}\cdots H_{i_{2s_N-1}i_o}(\star\star).$$

Due to symmetry, independence and zero mean assumption, each non oriented "edge" (ij) is seen an even number of times.

To each term in $(\star\star)$, we associate:

- a path $i_oi_1\cdots i_{2s_N-1}i_o$
- a trajectory $x(t), 0 \leq t \leq 2s_N$ starting at the origin and making $\pm$ steps. If at the instant t, the edge we see has been read for an odd number of times, then $+$ step $(1,1)$ and $-$ step $(1,-1)$ otherwise. This defines a Dyck path i.e. a trajectory in the positive quadrant of length $2s_N$ and ending at level 0.

Marked instants: right endpoint of an up edge.
Same as the classical proof of Wigner's theorem.

Paths

Given a trajectory $x(t)$, assign labels chosen amongst $\{1, \ldots, N\}$

- choose the origin i_o and vertices at marked instants,
- then "close" the edges by assigning vertices at non-marked instants.

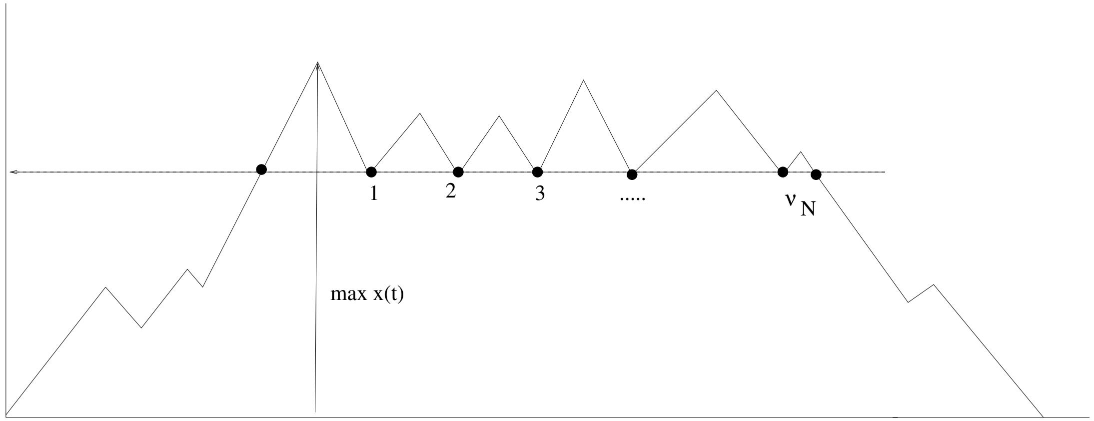
Wigner's regime: choose the marked vertices and origin pairwise distinct. No choice to close the edges.
Largest eigenvalue: $s_N \sim N^{2/3}$: repeat some marked vertices. This decreases the number of labels of a factor N but s_N^2 moments where a label occurs twice e.g.

Self intersection of type i: v occurs i times as a marked vertex.

Problem: which trajectories are typical? which paths are typical?

Typical paths and typical trajectories

Typical trajectories:



$$\max x(t) \sim \sqrt{s_N}, \quad \nu_N << \sqrt{s_N}.$$

Typical paths: The typical number of vertices of type i is $\left(\frac{s_N}{N}\right)^i N$.

In the scale $s_N \sim N^{2/3}$, there are self-intersections of type 3 at most.

Each edge is read at most twice in typical paths

This implies universality.

In the scale $N^{2/3}$ there are multiple choices to close the edges (GOE or GUE TW).

Random matrices with non symmetrically distributed entries

Let μ be a non symmetric distribution with compact support s.t. $\int x d\mu = 0$, $\int |x|^2 d\mu = \sigma^2$.

Conjecture: the limiting distribution of the largest eigenvalue is Tracy-Widom.

Theorem: Soshnikov, P. (2007)

$$\lambda_1 \leq 2\sigma + O(N^{-6/11}), \text{ with } p \to 1.$$

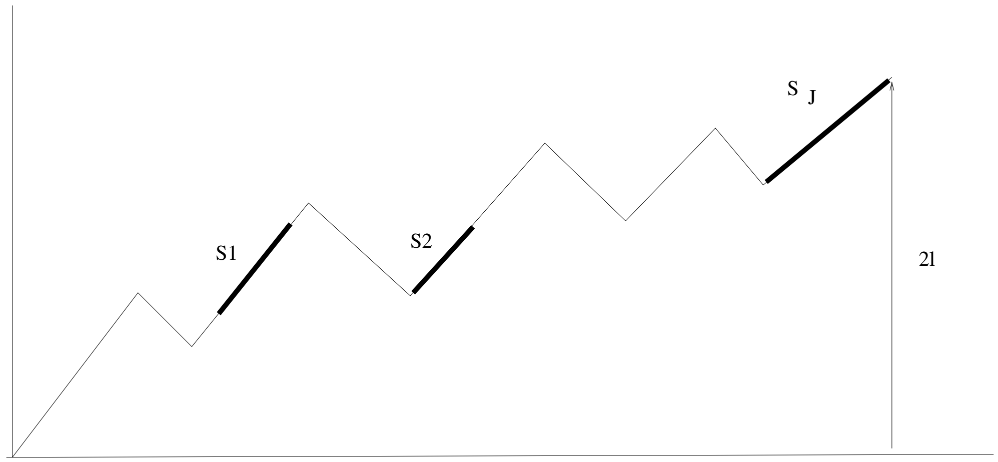
Improves a bound obtained earlier by Vu (2005)

$$\lambda_1 \leq 2\sigma + O(n^{-1/4} \ln n).$$

See also Füredi-Komlos (1981) and Guionnet-Zeitouni (2000)

Combinatorics

Paths with $2l$ edges read an odd number of times (at least three times).



Sequences $S_i, 1 \leq i \leq J$ where we read the last occurence of each odd edge. Erase these S_i's, we obtain (up to permutation) a path which is of the same kind as those occuring in the Wigner case.
Count the number of ways to choose the edges one repeats and the moments of time where one inserts the S_i's.
Show that paths with odd edges are negligible w.r.t. even paths.

Deformed Ensembles

Theorem : Féral, P.(2006)
H_N complex random matrix with i.i.d. entries with a symmetric distribution μ s.t.

- $\int |x|^2 d\mu(x) = \sigma^2$ and $\exists C > 0$ $\int |x|^{2k} d\mu(x) \leq (Ck)^k$, $\forall\ k$.
- $A_N = \frac{\pi_1}{N} J$ with $J_{ij} = 1, \forall i, j$: a specific rank one perturbation.

$$M_N = \frac{1}{\sqrt{N}} H_N + A_N.$$

First study: Komlos-Füredi Adjacency matrix of random graphs $\pi_1 \sim \sqrt{N}$.
Random matrix with non-centered entries.

Deformed GUE: H_N of the GUE.
Phase transition for the largest eigenvalue for any rank 1 perturbation A_N.

Universality results

Let λ_1 be the largest eigenvalue of $M_N = \frac{1}{\sqrt{N}} H_N + A_N$. Then,

$$\lim_{N\to\infty} P\left(N^{2/3}\left(\frac{\lambda_1}{\sigma} - 2\right) \leq x\right) = F_2^{TW}(x), \text{ if } \pi_1 < \sigma.$$

$$\lim_{N\to\infty} P\left(N^{2/3}\left(\frac{\lambda_1}{\sigma} - 2\right) \leq x\right) = F_3^{TW}(x), \text{ if } \pi_1 = \sigma.$$

$$\lim_{N\to\infty} P\left(\sigma^2(\pi_1) N^{1/2}\left(\frac{\lambda_1}{\sigma} - C(\pi_1)\right) \leq x\right) = P\left(\mathcal{N}(0, \sigma^2(\pi_1)) \leq (x)\right) \text{ if } \pi_1 > \sigma.$$

$$C(\pi_1) = \frac{\pi_1}{\sigma} + \frac{\sigma}{\pi_1} \quad \text{and} \quad \sigma^2(\pi_1) = \frac{\pi_1^2}{\pi_1^2 - \sigma^2}.$$

Idea of the proof:

typically the sequences of odd edges correspond to edges read only once.

Same typical paths as for standard Wigner random matrices, except that trajectories end at some level > 0.

Universality and unitary invariance

If $A_N = \text{diag}(\pi_1, 0, \ldots, 0)$ with $\pi_1 > \sigma$, then

$$\lambda_1(DGUE) = \pi_1 + \frac{\sigma^2}{\pi_1} + \xi N^{-1/2}, \text{ with } \xi \sim \mathcal{N}(0, \tilde{\sigma}^2(\pi_1)).$$

But consider the largest eigenvalue of

$$M_N = \frac{1}{\sqrt{N}} H_N + A_N$$

with H_N having entries with sub-Gaussian tails and A_N diagonal:
still true that $\lim_{N\to\infty} \lambda_1 = \pi_1 + \frac{\sigma^2}{\pi_1}$
but strongly suspect that the limiting fluctuations of λ_1 depend on μ.

Idea: the leading term in the asymptotic expansion of $\mathbb{E}\text{Tr}\left(\frac{M_N}{C(\pi_1)}\right)^{\sqrt{N}}$ is a function of the Laplace transform of μ.

Sample covariance matrices

λ_1 the largest eigenvalue of $\dfrac{1}{N}XX^*$, X: $N \times p$ random matrix (sub-Gaussian tail, symmetric distribution).

Set $C_N = (\frac{1}{\sqrt{N}} + \frac{1}{\sqrt{p}})^{-1/3}(\frac{1}{\sqrt{N}+\sqrt{p}})$, $u_+ = \sigma^2(1+\sqrt{\frac{p}{N}})^2$.

Theorem (2007) If $N, p \to \infty$ and $p/N \to \gamma \in [0, \infty]$

$$\lim_{N\to\infty} P\left(C_N\left(\lambda_1 - u_+\right) \le x\right) = F_2^{TW}(x).$$

Extends the result of Soshnikov: $\gamma = 1$ (based on computations for Wigner matrices).

$$\mathbb{E}\mathrm{Tr}\left[\frac{XX^*}{Nu_+}\right]^{s_N} \le \mathbb{E}\mathrm{Tr}\left[\frac{Y}{\sqrt{Nu_+}}\right]^{2s_N}$$

for some Wigner Hermitian random matrix Y.

Need to consider oriented edges.

Idea of the proof

Developping the trace $\sum_{i_o,\ldots,i_{s_N-1}} \sum_{j_o,\ldots,j_{s_N-1}} \mathbb{E}\left(X_{i_o j_o}\overline{X_{i_1 j_o}}\cdots X_{i_{s_N-1} j_{s_N-1}}\overline{X_{i_o j_{s_N-1}}}\right)$. We associate the sequence of oriented edges:

$$\binom{j_o}{i_o}\binom{j_o}{i_1}\binom{j_1}{i_1}\cdots\binom{j_{s_N-1}}{i_{s_N-1}}\binom{j_{s_N-1}}{i_o}$$

Define marked instants as before except that edges are oriented and read from bottom to top.

We still get a Dyck trajectory. What matters: number of odd and even marked instants (up steps) in Dyck paths. Indeed, p choices for labels instead of N.

Idea of the proof 2

Let $1 \leq k \leq s_N$.

$$\mathbf{N}(s_N, k) = \frac{1}{s_N} C_{s_N}^{k} C_{s_N}^{k-1}$$

Narayana number counts the number of Dyck trajectories of length $2s_N$ with k odd up steps.

Bai (1999) Connection with Marchenko-Pastur distribution

$$\sigma^{2l} \sum_{k=1}^{l} N(l,k)\gamma^k = \lim_{N \to \infty} \frac{1}{N} \mathbb{E}\mathrm{Tr} \left(\frac{XX^*}{N} \right)^l, \text{ if } \lim_{N \to \infty} \frac{p}{N} = \gamma.$$

Consider self-intersections on the top and bottom line separately.
Same statistics as for Wigner case.

Sample covariance matrices

- A diagonal covariance
  Under (slow) progress (w. D. Féral)

$$\frac{1}{N} X \Sigma X^*, \Sigma = \mathrm{diag}(\pi_1, 1, \ldots, 1)$$

  X $N \times p$ random matrix with i.i.d. complex entries (sub Gaussian tails).
  Universality is expected.

  Combinatorics: count the number of times a vertex occurs on the top row in typical paths.

- More general covariance Out of scope.

- Real sample covariance matrices No results for real Wishart matrices with non identity covariance.

Conclusion

- The method of moments: combinatorics highly depend on the structure of the deformation.
  For each "deformation" of a standard random matrix, new combinatorial approach.

- The resolvent approach: works only when the largest eigenvalue is well separated from the "bulk" of the spectrum.
  Silverstein (not published)
  Real sample covariance matrices with a spiked covariance matrix Σ ($Id +$ fixed rank)
  Gaussian fluctuations for the largest eigenvalues.

- What is the role of unitary invariance in Deformed Ensembles w.r.t. universality?