Global fluctuations for the β-Hermite and β-Laguerre ensembles

Ioana Dumitriu

joint work with **Ofer Zeitouni**

University of Washington (starting 09/06)

July 13, 2006

Outline.

1. β-ensembles: Hermite as generalized *GOE, GUE, GSE*, Laguerre as generalized *Wishart real, complex, quaternion.*

2. Semicircular laws and global fluctuations.

3. Sketch of proof for the Hermite case.

Gaussian orthogonal, unitary, and symplectic ensembles.

- $\mathbb{R}$, $\boldsymbol{\beta = 1}$: $A = \text{randn(n)}$, $S = (A + A')/2$;
- $\mathbb{C}$, $\boldsymbol{\beta = 2}$: $A = \text{randn(n)} + i * \text{randn(n)}$, $H = (A + A')/2$;
- $\mathbb{H}$, $\boldsymbol{\beta = 4}$: $A = \text{randn(n)} + i * \text{randn(n)} + j * \text{randn(n)} + k * \text{randn(n)}$,
  $D = (A + A')/2$.

Eigenvalue ($\Lambda := (\lambda_1, \ldots, \lambda_n)$) distributions:

$$\rho_{\boldsymbol{\beta}}(\Lambda) \propto \prod_{1 \le i < j \le n} |\lambda_i - \lambda_j|^{\boldsymbol{\beta}} \; e^{-\sum_{i=1}^{n} \lambda_i^2/2}$$

$$\equiv |\Delta(\Lambda)|^{\boldsymbol{\beta}} \; e^{-\sum_{i=1}^{n} \lambda_i^2/2} .$$

General β-Hermite (β-Gaussian) ensembles.

Defined through eigenvalue p.d.f. for $\boldsymbol{\beta} > 0$:

$$\propto \prod_{1\le i<j\le n} |\lambda_i - \lambda_j|^{\boldsymbol{\beta}} \quad e^{-\sum_{i=1}^{n} \lambda_i^2/2}$$

$$\equiv \quad |\Delta(\Lambda)|^{\boldsymbol{\beta}} \quad e^{-\sum_{i=1}^{n} \lambda_i^2/2} .$$

General β-Hermite (β-Gaussian) ensembles.

...or, equivalently, through their matrix model (with eigenvalue p.d.f. shown before):

$$H_\beta \sim \frac{1}{\sqrt{2}} \begin{pmatrix} N(0,2) & \chi_{(n-1)\beta} & & & \\ \chi_{(n-1)\beta} & N(0,2) & \chi_{(n-2)\beta} & & \\ & \ddots & \ddots & \ddots & \\ & & \chi_{2\beta} & N(0,2) & \chi_{\beta} \\ & & & \chi_{\beta} & N(0,2) \end{pmatrix}$$

(D. and Edelman, 2001)

Wishart real, complex, and quaternion ensembles.

- $\mathbb{R}$, $\boldsymbol{\beta} = \mathbf{1}$: $\quad A = \mathrm{randn(m,n)};\ W_1 = AA';$
- $\mathbb{C}$, $\boldsymbol{\beta} = \mathbf{2}$: $\quad A = \mathrm{randn(m,n)} + i{*}\mathrm{randn(m,n)};\ W_2 = AA';$
- $\mathbb{H}$, $\boldsymbol{\beta} = \mathbf{4}$: $\quad A = \mathrm{randn(m,n)} + i{*}\mathrm{randn(m,n)} + \ldots$
  $j{*}\mathrm{randn(m,n)} + k{*}\mathrm{randn(m,n)};\ W_4 = AA'.$

Eigenvalue ($\Lambda := (\lambda_1, \ldots, \lambda_m)$) distributions:

$$\rho_{\boldsymbol{\beta},n}(\Lambda) \propto \prod_{1\le i<j\le m} |\lambda_i - \lambda_j|^{\boldsymbol{\beta}} \prod_{i=1}^{m} \lambda_i^{\boldsymbol{\beta}(n-m+1)/2-1} \quad e^{-\sum_{i=1}^{m} \lambda_i/2}$$

$$\equiv |\Delta(\Lambda)|^{\boldsymbol{\beta}} \quad \det(\Lambda)^{\boldsymbol{\beta}(n-m+1)/2-1} \quad e^{-\sum_{i=1}^{m} \lambda_i/2} \; .$$

General β-Laguerre (β-Wishart) ensembles.

Defined through eigenvalue p.d.f. for $\beta > 0$, $a > \frac{\beta}{2}(n-1)$:

$$
\begin{aligned}
\propto \quad & \prod_{1\le i<j\le n} |\lambda_i - \lambda_j|^{\boldsymbol{\beta}} \prod_{i=1}^{n} \lambda_i^{a-\frac{\boldsymbol{\beta}}{2}(n-1)-1} \, e^{-\lambda_i/2} \\
\equiv \quad & |\Delta(\Lambda)|^{\boldsymbol{\beta}} \prod_{i=1}^{n} \lambda_i^{a-\frac{\boldsymbol{\beta}}{2}(n-1)-1} \, e^{-\lambda_i/2} \, .
\end{aligned}
$$

General β-Laguerre (β-Wishart) ensembles.

...or, equivalently, through their matrix model (with eigenvalue p.d.f. shown before):

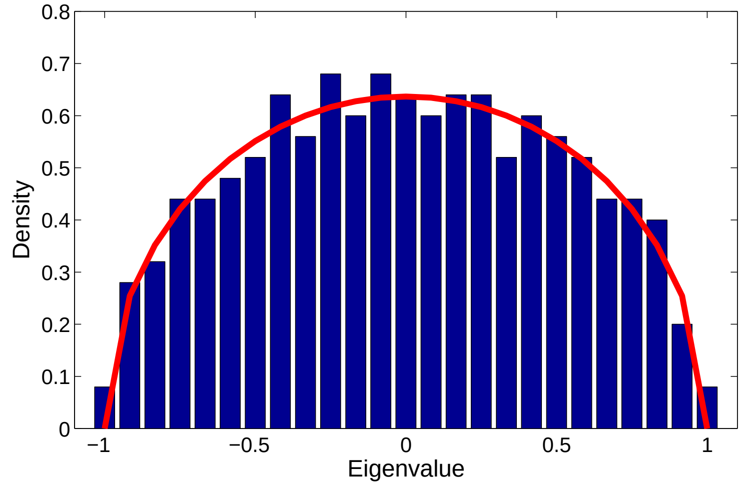
$$L_\beta = B_\beta B_\beta^T \ ,$$

where

$$B_\beta \sim \begin{pmatrix} \chi_{2a} & & & \\ \chi_{\beta(m-1)} & \chi_{2a-\beta} & & \\ & \ddots & \ddots & \\ & & \chi_\beta & \chi_{2a-\beta(m-1)} \end{pmatrix}$$

(D. and Edelman, 2001)

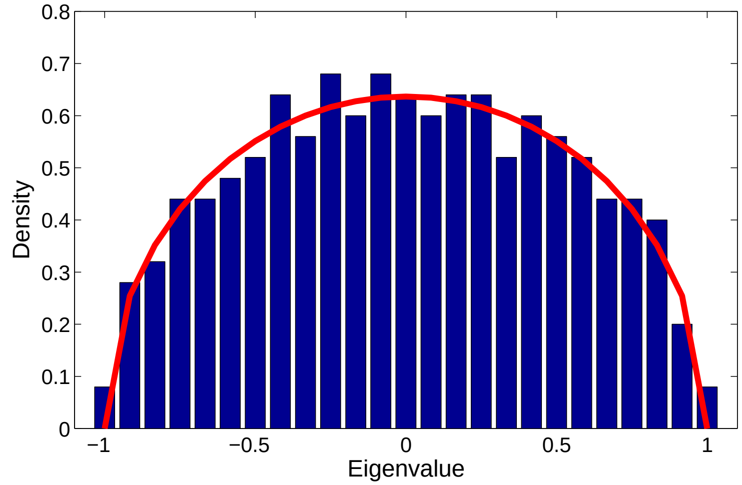
Semicircle Law.



Histogram of eig(S), where $n = 300$, $A =$randn(n), $S = (A + A')/(2\sqrt{2n})$

Works for all β-Hermite ensembles!

Semicircle Law.

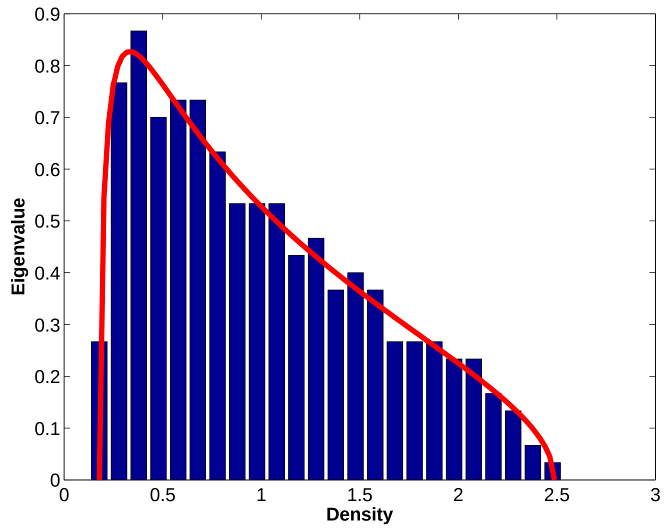


Histogram of eig(S), where $n = 300$, $A =$randn(n), $S = (A + A')/(2\sqrt{2n})$

Rigorously: Given any suitable function $h : [-1, 1] \longrightarrow \mathbb{R}$,

$$\frac{1}{n}\sum_{i=1}^{n} h(\lambda_i) \;\rightarrow\; \frac{2}{\pi}\int_{-1}^{1} h(x)\sqrt{1-x^2}\,dx$$

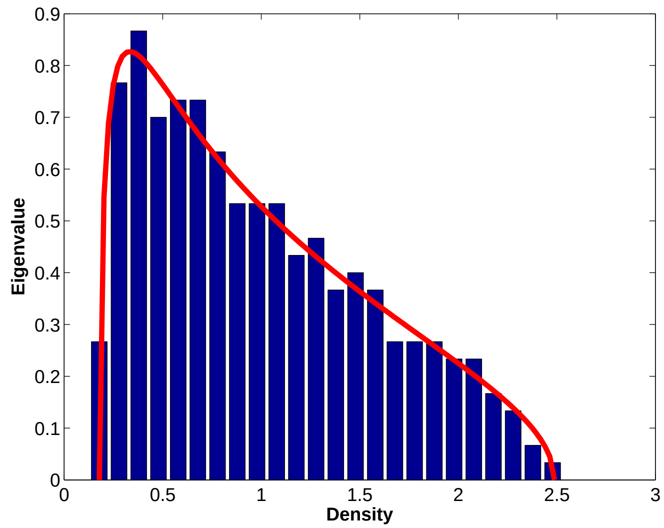
Marčenko-Pastur Law.



Histogram of eig(W), where $m = 300,\ n = 900,\ \gamma = 1/3,\ A =$randn$(m, n), W = AA'\ \sqrt{\gamma/m}$

Works for all β-Laguerre ensembles!

Marčenko-Pastur Law.



Histogram of eig(W), where $m = 300$, $n = 900$, $\gamma = 1/3$, A =randn(m, n), $W = AA' \sqrt{\gamma/m}$

Rigorously: Given any suitable function $h : [-1, 1] \longrightarrow \mathbb{R}$, let

$a = (1 - \sqrt{\gamma})^2$, $b = (1 + \sqrt{\gamma})^2$, then

$$\frac{1}{n}\sum_{i=1}^{n} h(\lambda_i) \;\rightarrow\; \frac{1}{2\pi\gamma}\int_a^b h(x)\frac{\sqrt{(b-x)(x-a)}}{x}\,dx$$

What about fluctuations from semicircle and Marčenko-Pastur laws?...

What about fluctuations from semicircle and Marčenko-Pastur laws?...

Fluctuations define Gaussian process.

Global fluctuations from semicircle and Marčenko-Pastur Laws.

CLT for Hermite. (Zeitouni and D., -)
Let h be a "nice" function, and λ_i, $i = \overline{1,n}$ the *scaled* eigenvalues of the β-Hermite ensemble. Then

$$\sum_{i=1}^{n} h(\lambda_i) - \frac{2n}{\pi}\int_{-1}^{1} h(x)\sqrt{1-x^2}\mathrm{d}x \quad \Longrightarrow$$

$$\Longrightarrow \quad N\left(\left(1-\frac{2}{\beta}\right)\mu(h), \frac{2}{\beta}\,\sigma^2(h)\right) \; .$$

"Nice" means **Lipschitz** on $[-1-\epsilon, 1+\epsilon]$ for some $\epsilon > 0$.

Works for **all** β.

Global fluctuations from semicircle and Marčenko-Pastur Laws.

CLT for Laguerre. (Zeitouni and D., -)
Let h be a "nice" function, and λ_i, $i = \overline{1,n}$ the *scaled* eigenvalues of the β-Laguerre ensemble (γ, a, b as before). Then

$$\sum_{i=1}^{n} h(\lambda_i) - \frac{n}{2\pi\gamma}\int_a^b h(x)\frac{\sqrt{(b-x)(x-a)}}{x}\mathrm{d}x \quad \Longrightarrow$$

$$\Longrightarrow \quad N\left(\left(1-\frac{2}{\beta}\right)\mu(h), \frac{2}{\beta}\,\sigma^2(h)\right) \; .$$

"Nice" means **Lipschitz** on $[-1-\epsilon, 1+\epsilon]$ for some $\epsilon > 0$.

Works for **all** β.

What was known before.

- **Johansson, 1998**

  CLT for Hermite ensembles, all β, h differentiable, extra conditions. *Also other even-degree polynomial potentials.*

What was known before.

- **Johansson, 1998**

  CLT for Hermite ensembles, all β, h differentiable, extra conditions. *Also other even-degree polynomial potentials.*

- **Bai & Silverstein, 2004**

  CLT for *Wishart-like* complex matrices, h analytic.

What was known before.

- **Johansson, 1998**

  CLT for Hermite ensembles, all β, h differentiable, extra conditions. *Also other even-degree polynomial potentials.*

- **Bai & Silverstein, 2004**

  CLT for *Wishart-like* complex matrices, h analytic.

- **Kusalik, Mingo, & Speicher, 2005**

  CLT for Wishart complex ($\beta = 2$), explicit calculation of covariance matrix, h polynomial. *Calculation of covariances for any finite number of independent matrices.*

What was known before.

- **Johansson, 1998**

  CLT for Hermite ensembles, all β, h differentiable, extra conditions. *Also other even-degree polynomial potentials.*

- **Bai & Silverstein, 2004**

  CLT for *Wishart-like* complex matrices, h analytic.

- **Kusalik, Mingo, & Speicher, 2005**

  CLT for Wishart complex ($\beta = 2$), explicit calculation of covariance matrix, h polynomial. *Calculation of covariances for any finite number of independent matrices.*

- **D. and Edelman, 2006**

  CLT for all β, Hermite and Laguerre, h polynomial. *Use of matrix models for all β.*

Outline of rest of the talk.

Use matrix methods, together with the "h polynomial" result of Edelman and D., to extend the class of h for which the Gaussian process is defined (in the talk, we only show this for Hermite).

1. Reduction to compactly supported functions h (Gershgorin).

2. Proof for a slightly different matrix model (Guionnet-Zeitouni).

3. Tying up the loose ends (Lidskii, Gershgorin, Chebyshev).

Outline of rest of the talk.

Use matrix methods, together with the "h polynomial" result of Edelman and D., to extend the class of h for which the Gaussian process is defined (in the talk, we only show this for Hermite).

1. Reduction to compactly supported functions h (Gershgorin).

2. Proof for a slightly different matrix model (Guionnet-Zeitouni).

3. Tying up the loose ends (Lidskii, Gershgorin, Chebyshev).

Sketch of idea.

Would like to approximate a "nice" function h by a sequence of polynomials p_n (on compact sets); remainder should be small with high probability.

Let $p_n \to h$ on $C \in \mathbb{R}$, $n||p_n - h|| \to 0$ (plus some other regularity). Then

$$\int_{-1}^{1} p_n(x)\sqrt{1-x^2}\mathrm{d}x \to \int_{-1}^{1} h(x)\sqrt{1-x^2}\mathrm{d}x \ ;$$

Need to show (e.g.) that for any C,

$$\Pr\left[\left| \sum_{i=1}^{n} h(\lambda_i) - \sum_{i=1}^{n} p_n(\lambda_i) \right| > C\right] \leq \frac{c}{C^2}\, ||p_n - h||\ .$$

Reduction to compactly supported functions.

We are essentially concerned with $\sum_{i=1}^{n} h(\lambda_i)$; we know (semicircle law) that

$$\frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i} \longrightarrow \mu_s \ ,$$

where $d\mu_s(x) = \frac{2}{\pi}\sqrt{1-x^2}$.

Enough to show that the probability of at least one of the eigenvalues being large (outside $[-1-\epsilon, 1+\epsilon]$, for fixed $\epsilon > 0$) decays *exponentially* in n.

To show this we use Gershgorin's theorem.

Reduction to compactly supported functions.

Recall the *scaled* matrix

$$H_{\beta} \sim \frac{1}{2\sqrt{n\beta}} \begin{pmatrix} N(0,2) & \chi_{(n-1)\beta} & & & \\ \chi_{(n-1)\beta} & N(0,2) & \chi_{(n-2)\beta} & & \\ & \ddots & \ddots & \ddots & \\ & & \chi_{2\beta} & N(0,2) & \chi_{\beta} \\ & & & \chi_{\beta} & N(0,2) \end{pmatrix}$$

All Gershgorin disks have centers distributed like $\tilde{c}\, N(0,1)/\sqrt{n}$, typical Gershgorin radius is distributed like $c\, \chi_{\alpha n}/\sqrt{n}$, $0 < \alpha < 1$.

But, for $\epsilon > 0$,

$$\Pr\left[(\tilde{c} |\, N(0,1)\,| \; + \; c\, \chi_{\alpha n}) \; > \; (1+\epsilon)\sqrt{n} \right] \leq c_1 e^{-c_2 n} \, ,$$

q.e.d.

A very useful result.

Theorem. (Guionnet-Zeitouni, 2000)
Let $A = (a_{ij})$ be a random matrix such that a_{ij} have distributions ν which satisfy the log-Sobolev inequality with uniform constant c, i.e., for all differentiable f,

$$\int f^2 \log \frac{f^2}{\int f^2 \mathrm{d}\nu} \mathrm{d}\nu \leq 2c \int |f'|^2 \mathrm{d}\nu \ ,$$

then for any Lipschitz function h, for any C,

$$\Pr \left[\left| \sum_{i=1}^{n} h(\lambda_i) - E\left[\sum_{i=1}^{n} h(\lambda_i) \right] \right| > C \right] \leq 2e^{-\frac{c_1}{h_{\mathcal{L}}} C^2} \ ,$$

where $h_{\mathcal{L}}$ is the Lipschitz constant for h.

Sounds perfect!

Sounds perfect!

except...

Sounds perfect!

except...

the χ function does not satisfy log-Sobolev constraints.

Sounds perfect!

except...

the χ function does not satisfy log-Sobolev constraints.

But the Gaussian does.

Fixup: approximate χ by Gaussian:

$$\chi_r \sim N(c_r, \frac{1}{\sqrt{2}}) + \quad small \ .$$

Outline of rest of the talk.

Use matrix methods, together with the "h polynomial" result of Edelman and D., to extend the class of h for which the Gaussian process is defined (in the talk, we only show this for Hermite).

1. Reduction to compactly supported functions h (Gershgorin).

2. Proof for a slightly different matrix model (Guionnet-Zeitouni).

3. Tying up the loose ends (Lidskii, Gershgorin, Chebyshev).

A slightly perturbed model.

The Guionnet-Zeitouni theorem works for

$$\tilde{H}_\beta \sim \frac{1}{2\sqrt{n\beta}} \begin{pmatrix} N(0,2) & N(c_{n-1},\frac{1}{2}) & & & \\ N(c_{(n-1)},\frac{1}{2}) & N(0,2) & N(c_{(n-2)},\frac{1}{2}) & & \\ & \ddots & \ddots & \ddots & \\ & & N(c_2,\frac{1}{2}) & N(0,2) & N(c_1,\frac{1}{2}) \\ & & & N(c_1,\frac{1}{2}) & N(0,2) \end{pmatrix}$$

Outline of rest of the talk.

Use matrix methods, together with the "h polynomial" result of Edelman and D., to extend the class of h for which the Gaussian process is defined (in the talk, we only show this for Hermite).

1. Reduction to compactly supported functions h (Gershgorin).

2. Proof for a slightly different matrix model (Guionnet-Zeitouni).

3. Tying up the loose ends (Lidskii, Gershgorin, Chebyshev).

Tying up loose ends.

Note that $H_{\text{diff}} = H_\beta - \tilde{H}_\beta$ looks, entry by entry, like

$$H_{\text{diff}} = O\left(\begin{pmatrix} 0 & \frac{1}{\sqrt{n(n-1)}} & & & \\ \frac{1}{\sqrt{n(n-1)}} & 0 & \frac{1}{\sqrt{n(n-2)}} & & \\ & \ddots & \ddots & \ddots & \\ & & \frac{1}{\sqrt{2n}} & 0 & \frac{1}{\sqrt{n}} \\ & & & \frac{1}{\sqrt{n}} & 0 \end{pmatrix}\right)$$

and the variables have uniformly bounded variance.

Tying up loose ends.

If λ_i are eigenvalues of H_β and $\tilde{\lambda}_i$ are eigenvalues of $\tilde{H}_\beta$, then for a Lipschitz h (with $h_{\mathcal{L}}$)

$$\left| \sum_{i=1}^{n} h(\lambda_i) - \sum_{i=1}^{n} h(\tilde{\lambda}_i) \right| \le h_{\mathcal{L}} \sum_{i=1}^{n} |\lambda_i - \tilde{\lambda}_i| \, .$$

Tying up loose ends.

If λ_i are eigenvalues of H_β and $\tilde{\lambda}_i$ are eigenvalues of $\tilde{H}_\beta$, then for a Lipschitz h (with $h_{\mathcal{L}}$)

$$\left|\sum_{i=1}^{n} h(\lambda_i) - \sum_{i=1}^{n} h(\tilde{\lambda}_i)\right| \;\le\; h_{\mathcal{L}} \sum_{i=1}^{n} |\lambda_i - \tilde{\lambda}_i| \;.$$

By **Lidskii's theorem**,

$$\sum_{i=1}^{n} |\lambda_i - \tilde{\lambda}_i| \le \sum_{i=1}^{n} |\lambda_{\mathrm{diff}_i}| \;,$$

where $\lambda_{\mathrm{diff}_i}$ are the eigenvalues of H_{diff}.

Tying up loose ends.

By **Gershgorin**,

$$\sum_{i=1}^{n} |\lambda_i - \tilde{\lambda}_i| \leq \sum_{i=1}^{n} O\left(\frac{1}{\sqrt{n}\sqrt{i}}\right) \ ,$$

and thus by **Chebyshev**,

$$\Pr\left[\left|\sum_{i=1}^{n} h(\lambda_i) - \sum_{i=1}^{n} h(\tilde{\lambda}_i)\right| \geq C\right] \ \leq \ h_{\mathcal{L}} \frac{c}{C^2} \ .$$

Wrap-up.

We obtained a (slightly) stronger CLT for β-Hermite and β-Laguerre ensembles, using a matrix-based analysis.

The largest class of functions we could define the Gaussian process on is **Lipschitz** on $[-1-\epsilon, 1+\epsilon]$, for some $\epsilon > 0$.

Natural technical condition: $\sigma^2(h) < \infty$.

Can regularity be relaxed to match this?...