The norm of polynomials in large random matrices

Camille Mâle,
École Normale Supérieure de Lyon,
Ph.D. Student under the direction of Alice Guionnet.

with a significant contribution by
Dimitri Shlyakhtenko.

ESI program Bialgebras in Free Probability, 11-21 April 2011.

A theorem of "strong asymptotic freeness"

The Gaussian Unitary Ensemble (GUE)

Definition

An $N \times N$ random matrix $X^{(N)}$ is a GUE matrix if $X^{(N)} = X^{(N)*}$ with entries $X^{(N)} = (X_{n,m})_{1 \leqslant n,m \leqslant N}$, where

$$\Big((X_{n,n})_{1 \leqslant n \leqslant N}, (\sqrt{2}\mathrm{Re}\ (X_{n,m}), \sqrt{2}\mathrm{Im}\ (X_{n,m})\ )_{1 \leqslant n < m \leqslant N} \Big)$$

is a centered Gaussian vector with covariance matrix $\frac{1}{N}\mathbf{1}_{N^2}$.

The Gaussian Unitary Ensemble (GUE)

Definition

An $N \times N$ random matrix $X^{(N)}$ is a GUE matrix if $X^{(N)} = X^{(N)*}$ with entries $X^{(N)} = (X_{n,m})_{1 \leqslant n,m \leqslant N}$, where

$$\Big((X_{n,n})_{1\leqslant n \leqslant N}, (\sqrt{2}\mathrm{Re}\ (X_{n,m}), \sqrt{2}\mathrm{Im}\ (X_{n,m})\ )_{1 \leqslant n < m \leqslant N} \Big)$$

is a centered Gaussian vector with covariance matrix $\frac{1}{N}\mathbf{1}_{N^2}$.

Density

$$\frac{1}{Z_N} \exp\Big(- \frac{1}{2N} \mathrm{Tr}\ X^2 \Big) \prod_{i \leqslant j} \mathrm{dRe} X_{i,j} \prod_{i<j} \mathrm{dIm} X_{i,j}.$$

$\Rightarrow$ standard Gaussian measure on $\Big(\mathrm{M}_N(\mathbb{C})_{Herm}, \langle A, B \rangle = \frac{1}{N}\mathrm{Tr}\ \big[AB \big] \Big)$.

The semicircle distribution

Definition

A non commutative random variable x in a *-probability space $(\mathcal{A}, .^*, \tau)$ has a semicircle distribution of $x = x^*$ and for any polynomial P,

$$\tau\big[P(x)\big] = \int P \mathrm{d}\sigma \text{ with } \mathrm{d}\sigma(x) = \frac{1}{2\pi}\sqrt{4 - x^2}\mathbf{1}_{|x|<2}\mathrm{d}x.$$

The semicircle distribution

Definition

A non commutative random variable x in a *-probability space $(\mathcal{A}, .^*, \tau)$ has a semicircle distribution of $x = x^*$ and for any polynomial P,

$$\tau\big[P(x)\big] = \int P \mathrm{d}\sigma \text{ with } \mathrm{d}\sigma(x) = \frac{1}{2\pi}\sqrt{4 - x^2}\mathbf{1}_{|x|<2}\mathrm{d}x.$$

Theorem: Wigner (58), Arnold (67)

if $X_N \rightsquigarrow$ GUE, then almost surely, for any polynomial P,

$$\frac{1}{N}\mathrm{Tr}\big[P(X_N)\big] = \frac{1}{N}\sum_{i=1}^{N} P\Big(\lambda_i(X_N)\Big) \underset{N\to\infty}{\longrightarrow} \int P\mathrm{d}\sigma = \tau\big[P(x)\big].$$

Voiculescu's asymptotic freeness theorem (1/2)

Consider in the *-probability space $\big(\mathrm{M}_N(\mathbb{C}), .^*, \tau_N := \frac{1}{N}\mathrm{Tr}\big)$

- $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ independent $N \times N$ GUE matrices,
- $\mathbf{Y}_N = (Y_1^{(N)}, \dots, Y_q^{(N)})$ $N \times N$ matrices, independent of $\mathbf{X}_N$.

Voiculescu's asymptotic freeness theorem (1/2)

Consider in the *-probability space $\left(\mathrm{M}_N(\mathbb{C}), .^*, \tau_N := \frac{1}{N}\mathrm{Tr}\right)$

- $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ independent $N \times N$ GUE matrices,
- $\mathbf{Y}_N = (Y_1^{(N)}, \dots, Y_q^{(N)})$ $N \times N$ matrices, independent of $\mathbf{X}_N$.

Consider in a *-probability space $(\mathcal{A}, .^*, \tau)$

- $\mathbf{x} = (x_1, \dots, x_p)$ free semicircular random variables,
- $\mathbf{y} = (y_1, \dots, y_q)$ free from $\mathbf{x}$.

Voiculescu's asymptotic freeness theorem (2/2)

Theorem: Voiculescu (91), Thorbjørnsen (99), Hiai and Petz (99)

Assume

- Almost surely, $\mathbf{Y}_N \xrightarrow{\mathcal{L}^{n.c.}} \mathbf{y}$ when $N \to \infty$ i.e. for every non commutative polynomial P,

$$\tau_N[P(\mathbf{Y}_N, \mathbf{Y}_N^*)] \underset{N\to\infty}{\longrightarrow} \tau[P(\mathbf{y}, \mathbf{y}^*)],$$

- Almost surely, for any $j = 1, \dots, q$, one has $\limsup_{N\to\infty} \| Y_j^{(N)} \| < \infty$.

Then almost surely $(\mathbf{X}_N, \mathbf{Y}_N) \xrightarrow{\mathcal{L}^{n.c.}} (\mathbf{x}, \mathbf{y})$ i.e. for every non commutative polynomial P,

$$\tau_N[P(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*)] \underset{N\to\infty}{\longrightarrow} \tau[P(\mathbf{x}, \mathbf{y}, \mathbf{y}^*)].$$

$\mathcal{C}^*$-Probability spaces

Definition

A $\mathcal{C}^*$-probability space $(\mathcal{A}, .^*, \tau, \|\cdot\|)$ consists of a *-probability space $(\mathcal{A}, .^*, \tau)$ and a norm $\|\cdot\|$ such that $(\mathcal{A}, .^*, \|\cdot\|)$ is a $\mathcal{C}^*$-algebra.

$\mathcal{C}^*$-Probability spaces

Definition

A $\mathcal{C}^*$-probability space $(\mathcal{A}, .^*, \tau, \|\cdot\|)$ consists of a *-probability space $(\mathcal{A}, .^*, \tau)$ and a norm $\|\cdot\|$ such that $(\mathcal{A}, .^*, \|\cdot\|)$ is a $\mathcal{C}^*$-algebra.

In $\big(\mathrm{M}_N(\mathbb{C}), .^*, \tau_N := \frac{1}{N}\mathrm{Tr}\big)$ we consider the operator norm:

$$\|A\| = \sqrt{\rho(A^*A)} \overset{A=A^*}{=} \rho(A).$$

$\mathcal{C}^*$-Probability spaces

Definition

A $\mathcal{C}^*$-probability space $(\mathcal{A}, .^*, \tau, \|\cdot\|)$ consists of a *-probability space $(\mathcal{A}, .^*, \tau)$ and a norm $\|\cdot\|$ such that $(\mathcal{A}, .^*, \|\cdot\|)$ is a $\mathcal{C}^*$-algebra.

In $\big(\mathrm{M}_N(\mathbb{C}), .^*, \tau_N := \frac{1}{N}\mathrm{Tr}\big)$ we consider the operator norm:

$$\|A\| = \sqrt{\rho(A^*A)} \overset{A=A^*}{=} \rho(A).$$

Proposition

If τ is faithful, i.e. $\tau[a^*a] = 0 \Rightarrow a = 0$, then

$$\|a\| = \lim_{k\to\infty} \Big(\tau\big[(a^*a)^k \big] \Big)^{\frac{1}{2k}}.$$

A strong asymptotic freeness theorem (1/2)

Consider in the $\mathcal{C}^*$-probability space $\big(\mathrm{M}_N(\mathbb{C}), .^*, \tau_N, \|\cdot\|\big)$

- $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ independent $N \times N$ GUE matrices,
- $\mathbf{Y}_N = (Y_1^{(N)}, \dots, Y_q^{(N)})$ $N \times N$ matrices, independent of $\mathbf{X}_N$.

Consider in a $\mathcal{C}^*$-probability space $(\mathcal{A}, .^*, \tau, \|\cdot\|)$ with **faithful** trace

- $\mathbf{x} = (x_1, \dots, x_p)$ free semicircular system,
- $\mathbf{y} = (y_1, \dots, y_q)$ free from $\mathbf{x}$.

A strong asymptotic freeness theorem (2/2)

Theorem: M. (11), preprint

Assume: Almost surely, for every non commutative polynomial P,

$$\tau_N\big[P(\mathbf{Y}_N, \mathbf{Y}_N^*)\big] \underset{N\to\infty}{\longrightarrow} \tau[P(\mathbf{y}, \mathbf{y}^*)],$$

$$\big\|P(\mathbf{Y}_N, \mathbf{Y}_N^*)\big\| \underset{N\to\infty}{\longrightarrow} \|P(\mathbf{y}, \mathbf{y}^*)\|.$$

Then, almost surely, for every non commutative polynomial P,

$$\tau_N\big[P(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*)\big] \underset{N\to\infty}{\longrightarrow} \tau[P(\mathbf{x}, \mathbf{y}, \mathbf{y}^*)],$$

$$\big\|P(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*)\big\| \underset{N\to\infty}{\longrightarrow} \|P(\mathbf{x}, \mathbf{y}, \mathbf{y}^*)\|.$$

Other results on strong asymptotic freeness

Cases where $\mathbf{Y}_N$ are zeros.

- Haagerup and Thorbjørnsen (05): pioneering works,
- Schultz (05): $\mathbf{X}_N \rightsquigarrow$ GOE, GSE,
- Capitaine and Donati-Martin (07): $\mathbf{X}_N \rightsquigarrow$ Wigner ensemble with symmetric law of entries and a concentration assumption; $\mathbf{X}_N \rightsquigarrow$ Wishart,
- Anderson (24 Mar 2011 on arXiv): $\mathbf{X}_N \rightsquigarrow$ Wigner ensemble with finite fourth moment.

The spectrum of large hermitian matrices

Corollary

Let $H_N = P(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*)$ a Hermitian matrix. Denote its empirical eigenvalue distribution

$$\mathcal{L}_{H_N} = \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i(H_N)}.$$

Asymptotic freeness: Almost surely, $\mathcal{L}_{H_N} \underset{N\to\infty}{\longrightarrow} \mathcal{L}_h$ the distribution of the self adjoint non commutative random variable $h = P(\mathbf{x}, \mathbf{y}, \mathbf{y}^*)$.
Strong asymptotic freeness: Almost surely, for every $\varepsilon > 0$, there exists $N_0 \geq 1$ such that for every $N \geq N_0$,

$$\mathrm{Sp}(\ H_N\ ) \subset \mathrm{Supp}\ (\ \mu\ ) + (-\varepsilon, \varepsilon).$$

Idea of the proof

The main steps

Haagerup and Thorbjørnsen's method:

1. A linearization trick,
2. Uniform control of matrix-valued Stieltjes transforms,
3. Concentration argument.

The main steps

Haagerup and Thorbjørnsen's method:

1. A linearization trick,
2. Uniform control of matrix-valued Stieltjes transforms,
3. Concentration argument.

In this proof

1. A linearization trick, unchanged,
2. Uniform control of matrix-valued Stieltjes transforms, based on an "asymptotic subordination property",
3. An intermediate inclusion of spectrum, by Shlyakhtenko,
4. Concentration argument, no significant changes.

An equivalent formulation

A linearization trick

In order to show: Almost surely, for every polynomial P one has
$\|P(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*)\| \underset{N\to\infty}{\longrightarrow} \|P(\mathbf{x}, \mathbf{y}, \mathbf{y}^*)\|$,
it is enough to show: Almost surely, for every self adjoint degree one polynomial L with coefficient in $\mathrm{M}_k(\mathbb{C})$, for any $\varepsilon > 0$, there exists $N_0 \geq 1$ such that for all $N \geq N_0$, one has

$$\mathrm{Sp}\big(\ L(\mathbf{X}_N, \mathbf{Y}_N, \mathbf{Y}_N^*)\ \big) \subset \mathrm{Sp}\big(\ L(\mathbf{x}, \mathbf{y}, \mathbf{y}^*)\ \big) + (-\varepsilon, \varepsilon).$$

Based on operator spaces techniques (Arveson's theorem and dilation of operators).

Matricial Stieltjes transforms and $\mathcal{R}$-transforms

Let $(\mathcal{A}, .^*, \tau, \|\cdot\|)$ be a $\mathcal{C}^*$-probability space. Consider z in $\mathrm{M}_k(\mathbb{C}) \otimes \mathcal{A}$.

Definitions

- The $\mathrm{M}_k(\mathbb{C})$-valued Stieltjes transform of z is

$$\begin{array}{rccc} G_z : & \mathrm{M}_k(\mathbb{C})^+ & \to & \mathrm{M}_k(\mathbb{C}) \\ & \Lambda & \mapsto & (\mathrm{id}_k \otimes \tau_N)\Big[(\Lambda \otimes \mathbf{1} - z)^{-1}\Big]. \end{array}$$

- The amalgamated $\mathcal{R}$-transform over $\mathrm{M}_k(\mathbb{C})$ of z is

$$\begin{array}{rccc} \mathcal{R}_z : & U & \to & \mathrm{M}_k(\mathbb{C}) \\ & \Lambda & \mapsto & G_z^{(-1)}(\Lambda) - \Lambda^{-1}. \end{array}$$

The subordination property

Let $\mathbf{x} = (x_1, \dots, x_p)$ and $\mathbf{y} = (y_1, \dots, y_q)$ be selfadjoint elements of $\mathcal{A}$ and let $\mathbf{a} = (a_1, \dots, a_p)$ and $\mathbf{b} = (b_1, \dots, b_q)$ be $k \times k$ Hermitian matrices. Define

$$s = \sum_{j=1}^{p} a_j \otimes x_j, \quad t = \sum_{j=1}^{q} b_j \otimes y_j.$$

Proposition

If the families $\mathbf{x}$ and $\mathbf{y}$ are free, then one has

$$G_{s+t}(\Lambda) = G_t\Big(\Lambda - \mathcal{R}_s\big(\, G_{s+t}(\Lambda)\,\big)\,\Big).$$

If $x_1, \dots, x_p$ are free semicircular n.c.r.v. then we get

$$\mathcal{R}_s : \Lambda \mapsto \sum_{j=1}^{p} a_j \Lambda a_j.$$

Stability under analytic perturbations

If one has

$$\begin{aligned} G_{s+t}(\Lambda) &= G_t\Big(\Lambda - \mathcal{R}_s\big(\, G_{s+t}(\Lambda)\,\big)\,\Big), \\ G(\Lambda) &= G_t\Big(\Lambda - \mathcal{R}_s\big(G(\Lambda)\,\big)\,\Big) + \Theta(\Lambda), \end{aligned}$$

where Θ is an analytic perturbation, then we get

$$\|G(\Lambda) - G_{s+t}(\Lambda)\| \leqslant \big(1 + c\, \|(\mathrm{Im}\ \Lambda)^{-1}\|^2\ \big)\ \|\Theta(\Lambda)\|.$$

An asymptotic subordination property

Let $\mathbf{X}_N = (X_1^{(N)}, \dots, X_p^{(N)})$ be independent GUE matrices, let $\mathbf{Y}_N = (Y_1^{(N)}, \dots, Y_q^{(N)})$ be deterministic Hermitian matrices and let $\mathbf{a} = (a_1, \dots, a_p)$ and $\mathbf{b} = (b_1, \dots, b_q)$ be $k \times k$ Hermitian matrices. Define

$$S_N = \sum_{j=1}^{p} a_j \otimes X_j^{(N)}, \quad T_N = \sum_{j=1}^{q} b_j \otimes Y_j^{(N)}.$$

Proposition

One has

$$G_{S_N+T_N}(\Lambda) = G_{T_N}\Big(\Lambda - \mathcal{R}_s\big(\, G_{S_N+T_N}(\Lambda) \, \big) \Big) + \Theta_N(\Lambda),$$

with Θ_N an analytic perturbation.

A first try

Let $\mathbf{x} = (x_1, \dots, x_p)$ be free semicircular n.c.r.v. and $\mathbf{y} = (y_1, \dots, y_q)$ the limit in law of $\mathbf{Y}_N$.

$$\begin{aligned} G_{s+t}(\Lambda) &= G_t\Big(\Lambda - \mathcal{R}_s\big(\ G_{s+t}(\Lambda)\ \big)\ \Big), \\ G_{S_N+T_N}(\Lambda) &= G_{T_N}\Big(\Lambda - \mathcal{R}_s\big(\ G_{S_N+T_N}(\Lambda)\ \big)\ \Big) + \Theta_N(\Lambda). \end{aligned}$$

$\Rightarrow$ we get an estimate of $\|G_{S_N+T_N}(\Lambda) - G_{s+t}(\Lambda)\|$ only if we can control $\|G_{T_N}(\Lambda) - G_t(\Lambda)\|$.
$\Rightarrow$ with the concentration machinery we get the Theorem, but with unsatisfactory assumptions on $\mathbf{Y}_N$ ...

An intermediate space

Put $\mathbf{x}$ and $\mathbf{Y}_N$ in a same $\mathcal{C}^*$-probability space, free from each other. Then

$$\begin{aligned} G_{s+T_N}(\Lambda) &= G_{T_N}\Big(\Lambda - \mathcal{R}_s\big(\, G_{s+T_N}(\Lambda)\,\big)\,\Big), \\ G_{S_N+T_N}(\Lambda) &= G_{T_N}\Big(\Lambda - \mathcal{R}_s\big(\, G_{S_N+T_N}(\Lambda)\,\big)\,\Big) + \Theta_N(\Lambda). \end{aligned}$$

$\Rightarrow$ we get an estimate of $\|G_{S_N+T_N}(\Lambda) - G_{s+T_N}(\Lambda)\|$ without any additionnal assumption on $\mathbf{Y}_N$.

An theorem about norm convergence

Theorem: by Shlyakhtenko, in an appendix of M. (11)

Let $\mathbf{Y}_N = (Y_1^{(N)}, \dots, Y_q^{(N)})$ and $\mathbf{y} = (y_1, \dots, y_q)$ be n.c.r.v. in a $\mathcal{C}^*$-probability space with faithful trace. Let $\mathbf{x} = (x_1, \dots, x_q)$ be free semicircular n.c.r.v. free from $\mathbf{Y}_N$ and $\mathbf{Y}$. Assume: for every non commutative polynomial P,

$$\begin{aligned} \tau_N\big[P(\mathbf{Y}_N, \mathbf{Y}_N^*)\big] &\underset{N\to\infty}{\longrightarrow} \tau[P(\mathbf{y}, \mathbf{y}^*)], \\ \big\|P(\mathbf{Y}_N, \mathbf{Y}_N^*)\big\| &\underset{N\to\infty}{\longrightarrow} \|P(\mathbf{y}, \mathbf{y}^*)\|, \end{aligned}$$

Then for every non commutative polynomial P,

$$\|P(\mathbf{x}, \mathbf{Y}_N, \mathbf{Y}_N^*)\| \underset{N\to\infty}{\longrightarrow} \|P(\mathbf{x}, \mathbf{y}, \mathbf{y}^*)\|.$$

An intermediate inclusion of spectrum

We get: for every self adjoint degree one polynomial L with coefficient in $\mathrm{M}_k(\mathbb{C})$, for any $\varepsilon > 0$, there exists $N_0 \geq 1$ such that for all $N \geq N_0$, one has

$$\mathrm{Sp}\big(\ L(\mathbf{x}, \mathbf{Y}_N, \mathbf{Y}_N^*)\ \big) \subset \mathrm{Sp}\big(\ L(\mathbf{x}, \mathbf{y}, \mathbf{y}^*)\ \big) + (-\varepsilon, \varepsilon).$$

$\Rightarrow$ Together with this estimate of $\| G_{S_N + T_N}(\Lambda) - G_{s+T_N}(\Lambda) \|$, the concentration machinery applies.

Thank you for your attention