A generalized Marchenko-Pastur Theorem

with Florent Benaych-Georges

Toronto, 2013

General setting

$$M^{(N)} = \sum_{p=1}^{K_N} X_p U_p^{(N)} U_p^{(N)*}$$

with

- $\lim_{N\to+\infty} \frac{K_N}{N} = \gamma > 0$ a.s.
- $(X_p, p \geq 1)$ real i.i.d. r.v., $X_p \sim \nu$;
- $\left(U_p^{(N)}, p \geq 1\right)$ i.i.d. with values in $\mathbb{C}^N$.

$\lambda_1^{(N)}, \ldots, \lambda_N^{(N)}$ e.v. of $M^{(N)}$; $\widehat{\mu}^{(N)} = \frac{1}{N}\sum_{i=1}^N \delta_{\lambda_i^{(N)}}$:

$$\widehat{\mu}^{(N)} \underset{N\to+\infty}{\longrightarrow} \mu^{(\infty)} \text{ ?}$$

Once upon a time : Marchenko-Pastur, 1967

Hypotheses

- k_N non random, $\frac{k_N}{N} \underset{N\to+\infty}{\longrightarrow} \gamma > 0$;
- $\mathbb{E}[\|U^{(N)}\|^4] < +\infty$; $\mathbb{E}[U^{(N)}(i)\bar{U}^{(N)}(j)] \sim \frac{\delta_{i,j}}{N}$;
  $\mathbb{E}[U^{(N)}(i)\bar{U}^{(N)}(j)U^{(N)}(k)\bar{U}^{(N)}(l)] \sim \frac{\delta_{i,j}\delta_{k,l}+\delta_{i,l}\delta_{j,k}}{N^2}$.

Theorem

$\widehat{\mu}^{(N)} \underset{N\to+\infty}{\overset{\mathbb{P}}{\longrightarrow}} \mu^{(\infty)}$ with $G_{\mu^{(\infty)}}(z) = \int \frac{1}{z-x}\mu^{(\infty)}(dx)$ s.t.

$$G_{\mu^{(\infty)}}(z) = \frac{1}{z - \gamma \int \frac{x}{1-xG_{\mu^{(\infty)}}(z)}\nu(dx)}$$

Free compound Poisson law (Speicher, 1998)

with rate $\gamma > 0$ and jump distribution ν :

$$
\begin{aligned}
CP^{\boxplus}(\gamma,\nu) &= \lim_{n\to+\infty} \left(\left(1-\frac{\gamma}{n}\right)\delta_0 + \frac{\gamma}{n}\nu \right)^{\boxplus n} \\
R_{CP^{\boxplus}(\gamma,\nu)}(z) &= \sum_{n\geq 1} z^{n-1} c_n^{\boxplus}\left(CP^{\boxplus}(\gamma,\nu)\right) = \gamma \int \frac{x}{1-zx}\nu(dx)
\end{aligned}
$$

Consequence

Since $G_{\mu^{(\infty)}}(z) = \frac{1}{z - R_{CP^{\boxplus}(\gamma,\nu)}(G_{\mu^{(\infty)}}(z))}$, then $\mu^{(\infty)} = CP^{\boxplus}(\gamma,\nu)$.

The Bercovici-Pata bijection for Infinitely Divisible laws

(Bercovici & Pata, Thorbjørnsen & Barndorff-Nielsen,...)

$$\Lambda : \begin{array}{rcl} ID^*(\mathbb{R}) & \to & ID^{\boxplus}(\mathbb{R}) \\ CP^*(\gamma,\nu) & \mapsto & CP^{\boxplus}(\gamma,\nu) \end{array}$$

with

$$CP^*(\gamma,\nu) = \lim_{n\to+\infty} \left(\left(1 - \frac{\gamma}{n}\right) \delta_0 + \frac{\gamma}{n}\nu \right)^{*n}$$

A consistent picture

$$CP^*(\gamma,\nu) \xrightarrow{\Lambda} CP^{\boxplus}(\gamma,\nu)$$

$$\sum_{p=1}^{P(\gamma)} X_p \qquad \sum_{p=1}^{k_N} X_p U_p^{(N)} U_p^{(N)*} \nearrow$$

$$N = 1 \qquad N \qquad N = +\infty$$

with $P(\gamma) \sim$ Poisson(γ), independent of $(X_p, p \geq 1)$.

Answer : replace k_N by $P(\gamma N)$.

Compound Poisson matrix model (Benaych-Georges,-, 2005)

$CP^{(N)}(\gamma, \nu)$: law of

$$\sum_{p=1}^{P(\gamma N)} X_p U_p^{(N)} U_p^{(N)*}$$

with

– $P(\gamma N) \sim$ Poisson(γN) ;

– $U_p^{(N)}$ uniformly distributed on $S^{N-1} = \left\{u \in \mathbb{C}^N / \|u\|_2 = 1\right\}$.

Properties of $CP^{(N)}(\gamma,\nu)$

– Nice matricial cumulants ;

– Nice Fourier transform :

$$\mathbb{E}\left[e^{\left(ia\sum_{p=1}^{P(\gamma)} X_p\right)}\right] \;=\; e^{\left(\gamma\int\left(e^{iya}-1\right)\nu(dy)\right)}$$

$$\mathbb{E}\left[e^{\left(i\mathrm{tr}A\sum_{p=1}^{P(\gamma N)} X_p U_p^{(N)} U_p^{(N)*}\right)}\right] \;=\; e^{\left(N\mathbb{E}\left[\gamma\int\left(e^{iyU^{(N)*}AU^{(N)}}-1\right)\nu(dy)\right]\right)}$$

– Can be extended to all infinitely divisible laws.

Heavy tailed covariance random matrices

(Belinschi, Dembo, Guionnet, 2009)

Hypotheses

- $\frac{k_N}{N} \longrightarrow \gamma$;
- $X_p = 1$, $\nu = \delta_1$;
- $U^{(N)}(k) = \frac{V_k}{a_N}$;
- $(V_k, k \geq 1)$ i.i.d. in the domain of attraction of an α-stable law.

Theorem

– $\widehat{\mu}^{(N)} \underset{N\to+\infty}{\longrightarrow} \mu^{(\infty)}(\alpha,\gamma)$ a.s.

– If $\gamma \in (0,1)$, then

$$\begin{aligned}
\mu^{(\infty)}(\alpha,\gamma) &= (1-\gamma)\delta_0 + \sigma_{\alpha,\gamma} \\
\frac{\mathrm{d}\sigma_{\alpha,\gamma}}{\mathrm{d}t}(t) &= -\frac{1}{\pi t}\Im\left(h_\alpha(Y_1(\sqrt{t}))\right) \\
z^\alpha Y_1(z) &= \frac{\gamma}{1+\gamma} C_\alpha g_\alpha(Y_2(z)) \\
z^\alpha Y_2(z) &= \frac{\gamma}{1+\gamma} C_\alpha g_\alpha(Y_1(z))
\end{aligned}$$

with $g_\alpha(y) = \int_0^{+\infty} t^{\frac{\alpha}{2}} e^{-yt^{\frac{\alpha}{2}}} \frac{e^{-t}}{t} dt$ and $h_\alpha(y) = \int_0^{+\infty} e^{-yt^{\frac{\alpha}{2}}} e^{-t} dt$.

A generalized Marchenko-Pastur theorem

(-,Benaych-Georges, 2012)

Hypotheses

- $\left(U^{(N)}(1), \ldots, U^{(N)}(N)\right)$ exchangeable ;
- $N^{\#\{i_1,\ldots,i_{2p}\}}\mathbb{E}\left[U^{(N)}(i_1)\cdots U^{(N)}(i_p)\overline{U^{(N)}(i_{p+1})\cdots U^{(N)}(i_{2p})}\right]$ bounded ;
- $N^k\mathbb{E}\left[|U^{(N)}(1)|^{2n_1}\cdots|U^{(N)}(k)|^{2n_k}\right] \underset{N\to+\infty}{\longrightarrow} \Gamma(n_1,\ldots,n_k) \leq C^{n_1+\cdots+n_k}$.

Theorem

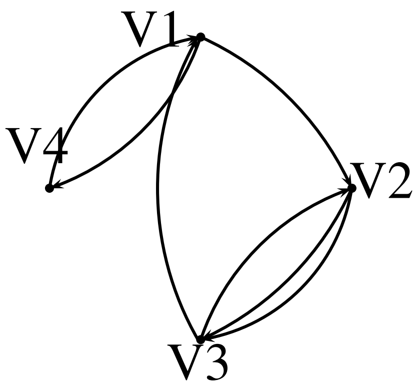
$$\widehat{\mu}^{(N)} \underset{N\to+\infty}{\longrightarrow} \mu^{(\infty)} =: \Lambda_{\Gamma}(CP^*(\gamma,\nu)) \text{ a.s.}$$

If $\forall k \geq 1$, $\nu(|x|^k) < +\infty$, then

$$\Lambda_{\Gamma}(CP^*(\gamma,\nu))(x^k) = \sum_{\pi\in\mathcal{P}(k)} f_{\Gamma}(\pi) c^*_{\pi}(CP^*(\gamma,\nu))$$

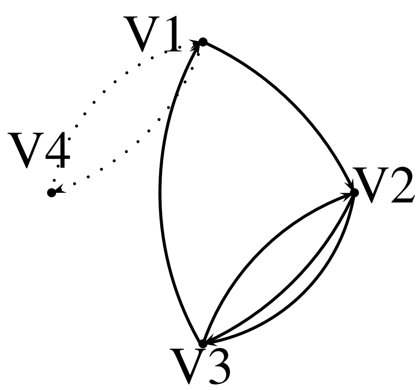
Computation of $f_\Gamma(\pi)$: the quotient graph

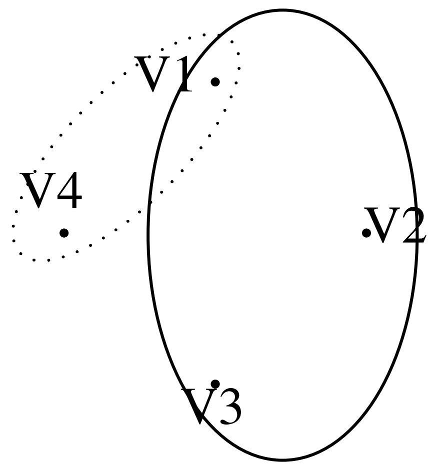
$\pi = \{\{1,6\},\{2,4\},\{3,5\},\{7\}\} =: \{V_1, V_2, V_3, V_4\}$



Computation of $f_\Gamma(\pi)$: the minimal colouring/hypertree

$\pi = \{\{1,6\},\{2,4\},\{3,5\},\{7\}\} =: \{V_1, V_2, V_3, V_4\}$





Computation of $f_\Gamma(\pi)$: the result

$\pi = \{\{1,6\},\{2,4\},\{3,5\},\{7\}\} =: \{V_1, V_2, V_3, V_4\}$

$f_\Gamma(\pi) = \varphi_{V_1}\varphi_{V_2}\varphi_{V_3}\varphi_{V_4}$:

- $\varphi_{V_1} = \lim_{N\to+\infty} \mathbb{E}\left[\left(\sum_{i=1}^N |U^{(N)}(i)|^2\right)^2\right]$;
- $\varphi_{V_2} = \lim_{N\to+\infty} \mathbb{E}\left[\sum_{i=1}^N |U^{(N)}(i)|^4\right]$;
- $\varphi_{V_3} = \lim_{N\to+\infty} \mathbb{E}\left[\sum_{i=1}^N |U^{(N)}(i)|^4\right]$;
- $\varphi_{V_4} = \lim_{N\to+\infty} \mathbb{E}\left[\sum_{i=1}^N |U^{(N)}(i)|^2\right]$.

Examples

Classical case

For $U^{(N)}$ uniformly distributed on the canonical basis, $\Lambda_\Gamma = Id$:

- $\forall n \geq 1$, $\sum_{i=1}^{N} |U^{(N)}(i)|^{2n} = 1$;
- $\forall \pi \in \mathcal{P}(k)$, $f_\Gamma(\pi) = 1$;
- $\int x^k \Lambda_\Gamma(CP^*(\gamma,\nu))(dx) = \sum_{\pi \in \mathcal{P}(k)} c_\pi^*(CP^*(\gamma,\nu)) = \int x^k CP^*(\gamma,\nu)(dx)$.

Not a surprise :

$$M^{(N)} = \begin{pmatrix} Y_1 & & \\ & \ddots & \\ & & Y_N \end{pmatrix} \text{ in law}$$

with $Y_1, \ldots, Y_N \sim CP^*(\gamma, \nu)$ i.i.d.

Free case

For $U^{(N)}$ uniformly distributed on S^{N-1}, or for $|U^{(N)}(j)| = \frac{1}{\sqrt{N}}$,
$\Lambda_\Gamma = \Lambda$:

- $\sum_{j=1}^{N} |U^{(N)}(j)|^{2n} \underset{N\to+\infty}{\longrightarrow} 0$ if $n > 1$;
- $f_\Gamma(\pi) = \mathbf{1}_{\pi\in\mathcal{NC}(k)}$;
- $\int x^k \Lambda_\Gamma(CP^*(\gamma,\nu))(dx) = \sum_{\pi\in\mathcal{NC}(k)} c^*_\pi(CP^*(\gamma,\nu)) = \int x^k CP^{\boxplus}(\gamma,\nu)(dx)$.

cf Marchenko-Pastur theorem.

Weak characterizations of $\mu^{(\infty)}$

$\Lambda_\Gamma(CP^*(\gamma,\nu))$ depends continuously of (γ,ν,Γ) :

1. If ν has no moment, let ν_C law of $X_1 \mathbf{1}_{|X_1|<c}$; then

$$\Lambda_\Gamma(CP^*(\gamma,\nu)) = \lim_{c\to+\infty} \Lambda_\Gamma(CP^*(\gamma,\nu_C))$$

2. Let $U^{(N)}$ as in Belinschi-Dembo-Guionnet theorem ; then

$$\begin{aligned} \widehat{\mu}^{(N)} \underset{N\to+\infty}{\longrightarrow} \mu^{(\infty)} &=: \mu^{(\infty)}(\alpha,\gamma,\nu) \\ &= \lim_{n\to+\infty} \Lambda_{\Gamma_n}(CP^*(\gamma,\nu)) \end{aligned}$$

   for appropriate $(\Gamma_n, n \geq 1)$.

Random measures

– $\mathcal{V}^{(N)} = \sum_{j=1}^{N} |U^{(N)}(j)|^2 \delta_{\frac{j}{N}}$ random measure on $(0,1]$.
– $\Delta_k = \left\{(x, \ldots, x) \in (0,1]^k\right\}$.

Then $\mathcal{V}^{(N)\otimes k}(\Delta_k) = \sum_{j=1}^{N} |U^{(N)}(j)|^{2k}$.

Hence f_Γ depends on $\lim_{N\to+\infty} \mathcal{V}^{(N)} := \mathcal{V}^{(\infty)}$.

Properties of $\mathcal{V}^{(\infty)}$

$\mathcal{V}^{(\infty)}$ is symmetrically distributed (cf Kallenberg) :

If $A_1, \dots, A_k \subset (0,1]$ are disjoints, then

$$\left(\mathcal{V}^{(\infty)}(A_1), \dots, \mathcal{V}^{(\infty)}(A_k)\right) \sim \mathit{Law}\,(\lambda(A_1), \dots, \lambda(A_k))$$

If $\lambda(A_1) = \cdots = \lambda(A_k)$, then $\left(\mathcal{V}^{(\infty)}(A_1), \dots, \mathcal{V}^{(\infty)}(A_k)\right)$ is an exchangeable random vector.

Moreover, if $U^{(N)}(1), \dots, U^{(N)}(N)$ are i.i.d., then $\mathcal{V}^{(\infty)}$ is a completely random measure (cf Kingman) : $\mathcal{V}^{(\infty)}(A_1), \dots, \mathcal{V}^{(\infty)}(A_k)$ are independent.

Characterization of $\mathcal{V}^{(\infty)}$ (Kallenberg, 1973)

$$\mathcal{V}^{(\infty)} = \eta\lambda_{(0,1]} + \sum_{i=1}^{+\infty} \beta_i \delta_{\tau_i}$$

with

- $\eta \geq 0$ r.v.
- $\beta_1 \geq \beta_2 \geq \cdots \geq 0$ r.v. such that $\sum_{i=1}^{+\infty} \beta_i < +\infty$ a.s.
- $\tau_1, \tau_2, \ldots$ i.i.d., uniformly distributed on $(0, 1]$;
- $(\eta, \beta_1, \beta_2, \ldots)$ independent of $(\tau_1, \tau_2, \ldots)$.

If $\mathcal{V}^{(\infty)}$ is a completely random measure, then

- $\eta = cte$;
- $\sum_{i=1}^{+\infty} \delta_{\beta_i}$ is a Poisson point process, with intensity ρ on $\mathbb{R}_+$ such that $\int y \wedge 1 \rho(dy) < +\infty$ (Campbell)

Examples

- Classical case : $\mathcal{V}^{(\infty)} = \delta_{\tau_1}$;
- Free case : $\mathcal{V}^{(\infty)} = \lambda_{(0,1]}$;
- Belinschi-Dembo-Guionnet case : $\mathcal{V}^{(\infty)}$ is a $\frac{\alpha}{2}$-stable random measure, with

$$\begin{aligned} \eta &= 0 \\ \rho(dy) &= \frac{\alpha}{2(1+\gamma)} \frac{dy}{y^{1+\frac{\alpha}{2}}} \end{aligned}$$

Combinatorial characterization of $\mu^{(\infty)}$

Let $\zeta(x) = \sum_{n\geq 1} \frac{x^n}{n!(n-1)!}$ and $M(z) = \sum_{n\geq 0} z^n \mu^{(\infty)}(x^n)$; then :

$$
\begin{aligned}
M(z) &= \int_0^{+\infty} e^{F(z,u)} e^{-u} du \\
G(z,x) &= \int_0^{+\infty} e^{F(z,u)} \zeta(zx \cdot u) \frac{e^{-u}}{u} du \\
F(z,x) &= \gamma \int_0^{+\infty} \int \mathbb{E}\Bigg[e^{zM(z)\cdot uy\eta} \sum_{k\geq 0} \frac{1}{k!} \sum_{i_1 \neq \cdots \neq i_k} \prod_{j=1}^{k} G(z, uy\beta_{i_j}) \\
&\qquad \times \Bigg(zx \cdot uy\eta + \sum_{i_0 \notin \{i_1,\ldots,i_k\}} \zeta(zx \cdot uy\beta_{i_0}) \Bigg) \Bigg] \frac{e^{-u}}{u} du \, \nu(dy)
\end{aligned}
$$

If $\mathcal{V}^{(\infty)}$ is a completely random measure with $\eta = 0$, then :

$$F(z,x) = \gamma \int_0^{+\infty} \int e^{\int G(z,uyv)\rho(dv)} \left(\int \zeta(zx \cdot uyv)\rho(dv) \right) \frac{e^{-u}}{u} du \, \nu(dy)$$

Examples

Classical case $\eta = 0,\ \beta_1 = 1,\ \beta_2 = \beta_3 = \cdots = 0$

$$\begin{aligned} F(z,x) &= \gamma \int_0^{+\infty} \int \zeta(xuyz) \frac{e^{-u}}{u} du\nu(dy) \\ &= \gamma \int \left(e^{yxz} - 1\right) \nu(dy) = F(zx) \end{aligned}$$

$$M(z) = \sum_{k\geq 0} z^k \mu^{(\infty)}(x^k) = \int_0^{+\infty} e^{F(zu)} e^{-u} du$$

$$\begin{aligned} \int e^{zx} \mu^{(\infty)}(dx) &= \sum_{k\geq 0} \frac{z^k}{k!} \mu^{(\infty)}(x^k) \\ &= \exp(F(z)) \\ &= \exp\left(\gamma \int \left(e^{yz} - 1\right) \nu(dy)\right) \end{aligned}$$

Hence $\mu^{(\infty)} = CP^*(\gamma, \nu)$.

Free case $\eta = 1,\ \beta_1 = \beta_2 = \beta_3 = \cdots = 0$

$$\begin{aligned} F(z,x) &= \gamma \int_0^{+\infty} \int xzuy\, e^{uyzM(z)} \frac{e^{-u}}{u} du\nu(dy) \\ &= xz\,\gamma \int \frac{y}{1 - yzM(z)} \nu(dy) = xzR_{CP^{\boxplus}(\gamma,\nu)}(zM(z)) \end{aligned}$$

$$M(z) = \int e^{F(z,u)} e^{-u} du = \frac{1}{1 - zR_{CP^{\boxplus}(\gamma,\nu)}(zM(z))}$$

$$G_{\mu^{(\infty)}}(z) = \frac{1}{z} M\left(\frac{1}{z}\right) = \frac{1}{z - R_{CP^{\boxplus}(\gamma,\nu)w}(G_{\mu^{(\infty)}}(z))}$$

Hence $\mu^{(\infty)} = CP^{\boxplus}(\gamma, \nu)$.

Belinschi-Dembo-Guionnet case

$\nu = \delta_1$, $\eta = 0$, $\sum_{i=1}^{+\infty} \delta_{\beta_i}$ PPP with intensity $\rho(dy) = \frac{\alpha}{2(1+\gamma)} \frac{dy}{y^{1+\frac{\alpha}{2}}}$

$$\begin{aligned} F(z,x) &= \gamma \int_0^{+\infty} \left(\int \zeta(xzuy)\rho(dy) \right) \exp \left(\int G(z,uy)\rho(dy) \right) \frac{e^{-u}}{u} du \\ G(z,x) &= \int_0^{+\infty} e^{F(z,u)} \zeta(xzu) \frac{e^{-u}}{u} du \end{aligned}$$

Since $\int \zeta(zy)\rho(dy) = \frac{1}{1+\gamma} C_\alpha z^{\frac{\alpha}{2}}$, then

$$
\begin{aligned}
X_1(z,x) &= x^{-\frac{\alpha}{2}} F(z,x) \\
&= \frac{\gamma}{1+\gamma} C_\alpha z^{\frac{\alpha}{2}} \int u^{\frac{\alpha}{2}} e^{X_2(z,u) u^{\frac{\alpha}{2}}} \frac{e^{-u}}{u} du = X_1(z) \\
X_2(z,x) &= x^{-\frac{\alpha}{2}} \int G(z,xy)\rho(dy) \\
&= \frac{1}{1+\gamma} C_\alpha z^{\frac{\alpha}{2}} \int u^{\frac{\alpha}{2}} e^{X_1(z,u) u^{\frac{\alpha}{2}}} \frac{e^{-u}}{u} du = X_2(z)
\end{aligned}
$$

If $Y_{1,2}(z) = -X_{1,2}(1/z^2)$, then

$$\begin{aligned} Y_1(z)z^\alpha &= \frac{\gamma}{1+\gamma} C_\alpha g_\alpha(Y_2(z)) \\ Y_2(z)z^\alpha &= \frac{1}{1+\gamma} C_\alpha g_\alpha(Y_1(z)) \end{aligned}$$

with $g_\alpha(y) = \int_0^{+\infty} t^{\frac{\alpha}{2}} e^{-yt^{\frac{\alpha}{2}}} \frac{e^{-t}}{t} dt$. Therefore :

$$\begin{aligned} G_{\mu(\infty)}(z) &= \frac{1}{z} M\left(\frac{1}{z}\right) = \frac{1}{z}\int_0^{+\infty} e^{F(1/z,u)} e^{-u} du \\ &= \frac{1}{z}\int_0^{+\infty} e^{X_1(1/z)u^{\frac{\alpha}{2}}} e^{-u} du \\ &= \frac{1}{z} h_\alpha\left(Y_1(\sqrt{z})\right) \end{aligned}$$

with $h_\alpha(y) = \int_0^{+\infty} e^{-yt^{\frac{\alpha}{2}}} e^{-t} dt$.

References

– V.A. Marchenko, L.A. Pastur *Distribution of eigenvalues for some sets of random matrices*. Mat. Sb. (N.S.), 72(114) :4
– F. Benaych-Georges *Classical and free infinitely divisible distributions and random matrices*. Ann. Probab. Vol. 33 (2005)
– T. Cabanal-Duvillard *A matrix representation of the Bercovici-Pata bijection*. Electron. J. Probab. 10 (2005)
– S.T. Belinschi, A. Dembo, A. Guionnet *Spectral measure of heavy tailed band and covariance random matrices*. Comm. Math. Phys. 289 (2009)
– F. Benaych-Georges, T. Cabanal-Duvillard *Marchenko-Pastur theorem and Bercovici-Pata bijections for heavy-tailed or localized vectors*. Lat. Am. J. Probab. Math. Stat. 9 (2012)
– C. Male *The distributions of traffics and their free product*. ArXiV (2011)