Outliers in the Spectrum of Spiked Deformations of Unitarily Invariant Random Matrices

Random Matrices and their Applications

Maxime FEVRIER, joint work with S. BELINSCHI, H. BERCOVICI and M. CAPITAINE.

Université Paris-Sud

Tuesday, October 9th 2012

Historical overview

- Definition of a spiked population model (sample covariance setting, Johnstone 2001):

Historical overview

- Definition of a spiked population model (sample covariance setting, Johnstone 2001):
  $X \, \mathrm{Diag}(\underbrace{\theta_1, \ldots, \theta_1}_{k_1}, \ldots, \underbrace{\theta_J, \ldots, \theta_J}_{k_J}, 1, \ldots, 1) X^*$.

Historical overview

- Definition of a spiked population model (sample covariance setting, Johnstone 2001): $X \operatorname{Diag}(\underbrace{\theta_1, \ldots, \theta_1}_{k_1}, \ldots, \underbrace{\theta_J, \ldots, \theta_J}_{k_J}, 1, \ldots, 1) X^*$.
- Results in the sample covariance setting: from Baik-Ben Arous-Péché (2005) to Bai-Yao (2012).

Historical overview - 2

- Definition of an additive analogue (Péché 2006):
  $W + \text{Diag}(\underbrace{\theta_1, \ldots, \theta_1}_{k_1}, \ldots, \underbrace{\theta_J, \ldots, \theta_J}_{k_J}, 0, \ldots, 0).$

Historical overview - 2

- Definition of an additive analogue (Péché 2006):
  $$W + \text{Diag}(\underbrace{\theta_1, \ldots, \theta_1}_{k_1}, \ldots, \underbrace{\theta_J, \ldots, \theta_J}_{k_J}, 0, \ldots, 0).$$
- Results for the additive analogue: from Péché (2006) to Renfrew-Soshnikov (preprint 2012).

Two works

Largest eigenvalues of finite rank perturbations of unitarily invariant random matrices.

Theorem (Benaych-Georges and Nadakuditi 2009)

Almost surely,

$$\lambda_j \to_{N\to+\infty} \begin{cases} G_\nu^{-1}(1/\theta_j) & \text{if } \theta_j > 1/\lim_{z\downarrow b} G_\nu(z), \\ b & \text{otherwise,} \end{cases}$$

while for each fixed $j > r$, almost surely, $\lambda_j \to_{N\to+\infty} b$. Here,

$$G_\nu : \mathbb{C} \setminus \mathrm{supp}(\nu) \to \mathbb{C}, \quad G_\nu(z) = \int_{\mathbb{R}} \frac{d\nu(t)}{z-t},$$

is the Cauchy-Stieltjes transform of the limit distribution ν, and b is the maximum of its support.

Two works

Eigenvalues of full rank perturbations of Wigner matrices.

Theorem (Capitaine-Donati-Martin-Féral and F. 2010)

Let $H(z) := z + \sigma^2 G_\mu(z)$, then there are k_j eigenvalues converging almost surely to $H(\theta_j)$ iff $H'(\theta_j) > 0$, where μ is the limit distribution of the perturbation, and σ^2 is the variance of the entries of the Wigner matrix.

Model

$$X_N = \qquad U_N^* B_N U_N,$$

- $B_N = \text{Diag}(\beta_1^{(N)}, \ldots, \beta_N^{(N)})$,
- U_N is a random $N \times N$ unitary matrix distributed according to Haar measure.

Model

$$X_N = A_N + U_N^* B_N U_N,$$

- $A_N = \text{Diag}(\underbrace{\theta_1, \ldots, \theta_1}_{k_1}, \ldots, \underbrace{\theta_J, \ldots, \theta_J}_{k_J}, \alpha_1^{(N)}, \ldots, \alpha_{N-r}^{(N)}),$
- $B_N = \text{Diag}(\beta_1^{(N)}, \ldots, \beta_N^{(N)}),$
- U_N is a random $N \times N$ unitary matrix distributed according to Haar measure.

Model

$$X_N = A_N + U_N^* B_N U_N,$$

- $A_N = \text{Diag}(\underbrace{\theta_1, \ldots, \theta_1}_{k_1}, \ldots, \underbrace{\theta_J, \ldots, \theta_J}_{k_J}, \alpha_1^{(N)}, \ldots, \alpha_{N-r}^{(N)}),$
- $B_N = \text{Diag}(\beta_1^{(N)}, \ldots, \beta_N^{(N)}),$
- U_N is a random $N \times N$ unitary matrix distributed according to Haar measure.

Question: Spectrum of $X_N = A_N + U_N^* B_N U_N$?

Assumptions

- $B_N = \mathrm{Diag}(\beta_1^{(N)}, \ldots, \beta_N^{(N)})$;

$$\mu_{B_N} := \frac{1}{N}\sum_{i=1}^{N} \delta_{\beta_i^{(N)}} \Rightarrow \nu \in \mathcal{P}_c(\mathbb{R}),$$

$$\max_{1\leq j\leq N} \operatorname{dist}(\beta_j^{(N)}, \operatorname{supp}(\nu)) \to_{N\to\infty} 0.$$

Assumptions

- $B_N = \text{Diag}(\beta_1^{(N)}, \dots, \beta_N^{(N)})$;

$$\mu_{B_N} := \frac{1}{N}\sum_{i=1}^{N} \delta_{\beta_i^{(N)}} \Rightarrow \nu \in \mathcal{P}_c(\mathbb{R}),$$

$$\max_{1\le j\le N} \mathrm{dist}(\beta_j^{(N)}, \mathrm{supp}(\nu)) \to_{N\to\infty} 0.$$

- $A_N = \text{Diag}(\underbrace{\theta_1, \dots, \theta_1}_{k_1}, \dots, \underbrace{\theta_J, \dots, \theta_J}_{k_J}, \alpha_1^{(N)}, \dots, \alpha_{N-r}^{(N)})$;

$$\mu_{A_N} := \frac{1}{N}\sum_{i=1}^{N} \delta_{\lambda_i(A_N)} \Rightarrow \mu \in \mathcal{P}_c(\mathbb{R}),$$

$$\max_{1\le j\le N-r} \mathrm{dist}(\alpha_j^{(N)}, \mathrm{supp}(\mu)) \to_{N\to\infty} 0,$$

$\theta_j \notin \mathrm{supp}(\mu)$(the so-called spikes).

Global behaviour

We will use the usual notation:

$$\mu_{X_N} := \frac{1}{N} \sum_{\lambda \in \mathrm{sp}(X_N)} \delta_\lambda .$$

Asymptotic freeness (Voiculescu 91, Speicher 93)

Under these assumptions,

$$\mu_{X_N} \underset{N\to+\infty}{\overset{a.s.}{\Rightarrow}} \mu \boxplus \nu .$$

Global behaviour

We will use the usual notation:

$$\mu_{X_N} := \frac{1}{N} \sum_{\lambda \in \mathrm{sp}(X_N)} \delta_\lambda .$$

Asymptotic freeness (Voiculescu 91, Speicher 93)

Under these assumptions,

$$\mu_{X_N} \underset{N\to+\infty}{\overset{a.s.}{\Longrightarrow}} \mu \boxplus \nu .$$

What is this $\boxplus$ operation?

Free convolution of measures

Given $\tau \in \mathcal{P}_c(\mathbb{R})$, one defines:

Stieltjes transform

$$G_\tau(z) = \int_{\mathbb{R}} \frac{d\tau(t)}{z - t}, \quad z \notin \mathbb{R}.$$

Free convolution of measures

Given $\tau \in \mathcal{P}_c(\mathbb{R})$, one defines:

Stieltjes transform

$$G_\tau(z) = \int_{\mathbb{R}} \frac{d\tau(t)}{z-t}, \quad z \notin \mathbb{R}.$$

R-transform

$$R_\tau(z) = G_\tau^{-1}(z) - \frac{1}{z}.$$

Free convolution of measures

Given $\tau \in \mathcal{P}_c(\mathbb{R})$, one defines:

Stieltjes transform

$$G_\tau(z) = \int_{\mathbb{R}} \frac{d\tau(t)}{z - t}, \quad z \notin \mathbb{R}.$$

R-transform

$$R_\tau(z) = G_\tau^{-1}(z) - \frac{1}{z}.$$

Definition

One calls free convolution of μ and ν the probability measure $\mu \boxplus \nu \in \mathcal{P}_c(\mathbb{R})$ characterized by:

$$R_{\mu \boxplus \nu}(z) = R_\mu(z) + R_\nu(z).$$

Subordination

Theorem (Voiculescu 93, Biane 98)

There is a unique analytic map $\omega : \mathbb{C}^+ \to \mathbb{C}^+$ such that:

$$\forall z \in \mathbb{C}^+, \ \ G_{\mu \boxplus \nu}(z) = G_\mu(\omega(z)).$$

Subordination

Theorem (Voiculescu 93, Biane 98)

There is a unique analytic map $\omega : \mathbb{C}^+ \to \mathbb{C}^+$ such that:

$$\forall z \in \mathbb{C}^+, \ \ G_{\mu \boxplus \nu}(z) = G_\mu(\omega(z)).$$

Lemma

The map ω has an extension to $\mathbb{C}$ so that:

(a) ω is continuous on $\mathbb{C}^+ \cup \mathbb{R}$;

(b) $\omega(\{\infty\} \cup \mathbb{R} \setminus \mathrm{supp}(\mu \boxplus \nu)) \subseteq \{\infty\} \cup \mathbb{R} \setminus \mathrm{supp}(\mu)$;

(c) $\forall z \in \mathbb{C} \setminus \mathbb{R}, \overline{\omega(z)} = \omega(\overline{z})$;

(d) ω is meromorphic on $\mathbb{C} \setminus \mathrm{supp}(\mu \boxplus \nu)$.

A definition

Definition

For each $j \in \{1, \dots, J\}$, define O_j the set of solutions in $\mathbb{R} \setminus \mathrm{supp}(\mu \boxplus \nu)$ of the equation

$$\omega(\rho) = \theta_j, \tag{1}$$

and

$$O = \bigcup_{1 \leq j \leq J} O_j.$$

Question

Where are precisely located the eigenvalues of
$X_N = A_N + U_N^* B_N U_N$?

Question

Where are precisely located the eigenvalues of
$X_N = A_N + U_N^* B_N U_N$?

- Are they all contained in a neighborhood of the support of $\mu \boxplus \nu$?

Question

Where are precisely located the eigenvalues of
$X_N = A_N + U_N^* B_N U_N$?

- Are they all contained in a neighborhood of the support of $\mu \boxplus \nu$?
- Or are some of them lying outside of $\mathrm{supp}(\mu \boxplus \nu)$ (the so-called outliers)?

Question

Where are precisely located the eigenvalues of $X_N = A_N + U_N^* B_N U_N$?

- Are they all contained in a neighborhood of the support of $\mu \boxplus \nu$?
- Or are some of them lying outside of $\mathrm{supp}(\mu \boxplus \nu)$ (the so-called outliers)?

Theorem (Collins-Male 2011)

If $r = 0$ (no spikes), then almost surely,

$$\forall \eta > 0, \exists N_0 \in \mathbb{N}, \forall N \geq N_0, \mathrm{sp}(X_N) \subseteq K_\eta,$$

where $K_\eta := \{x \in \mathbb{R} \mid d(x, \mathrm{supp}(\mu \boxplus \nu)) \leq \eta\}$.

Main result

In the general case, one proves:

Theorem

The following results hold almost surely:

- for each $\rho \in O_j$, for all small enough $\varepsilon > 0$, for all large enough N,

$$\mathrm{card}\{\mathrm{sp}(X_N) \bigcap]\rho - \epsilon; \rho + \epsilon[\} = k_j;$$

- for almost all $\eta > 0$, for all small enough $\varepsilon > 0$, for large enough N,

$$\mathrm{sp}(X_N) \bigcap \mathbb{C} \setminus K_\eta \subset \bigcup_{\rho \in O \bigcap \mathbb{C} \setminus K_\eta}]\rho - \epsilon; \rho + \epsilon[.$$

Generalization

Remark

Actually, our result holds for

$$\tilde{X_N} = \tilde{A_N} + \tilde{B_N},$$

where $\tilde{A_N}$ and $\tilde{B_N}$ are independent random Hermitian matrices, provided the distribution of $\tilde{B_N}$ is invariant by conjugation by unitary matrices.

Comments

In the particular case of a finite rank deformation A_N, one recovers the result of Benaych-Georges and Nadakuditi (BGN 2009) on the convergence of the largest eigenvalues:

Theorem (Benaych-Georges and Nadakuditi 2009)

Almost surely,

$$\lambda_j \to_{N \to +\infty} \begin{cases} G_\nu^{-1}(1/\theta_j) & \text{if } \theta_j > 1/\lim_{z \downarrow b} G_\nu(z), \\ b & \text{otherwise,} \end{cases}$$

while for each fixed $j > r$, almost surely, $\lambda_j \to_{N \to +\infty} b$.

Comments

In the particular case of a finite rank deformation A_N, one recovers the result of Benaych-Georges and Nadakuditi (BGN 2009) on the convergence of the largest eigenvalues:

Theorem (Benaych-Georges and Nadakuditi 2009)

Almost surely,

$$\lambda_j \to_{N \to +\infty} \begin{cases} G_\nu^{-1}(1/\theta_j) & \text{if } \theta_j > 1/\lim_{z \downarrow b} G_\nu(z), \\ b & \text{otherwise,} \end{cases}$$

while for each fixed $j > r$, almost surely, $\lambda_j \to_{N \to +\infty} b$.

Indeed, in that case, $\mu = \delta_0$ and $\omega(z) = \frac{1}{G_\nu(z)}$.

Comments

In the case of a full rank deformation of a GUE, one recovers the result of Capitaine, Donati-Martin, Féral and F. (CDFF 2010).

Theorem (Capitaine-Donati-Martin-Féral and F. 2010)

Let $H(z) := z + \sigma^2 G_\mu(z)$, then there are k_j eigenvalues converging almost surely to $H(\theta_j)$ iff $H'(\theta_j) > 0$.

Comments

In the case of a full rank deformation of a GUE, one recovers the result of Capitaine, Donati-Martin, Féral and F. (CDFF 2010).

Theorem (Capitaine-Donati-Martin-Féral and F. 2010)

Let $H(z) := z + \sigma^2 G_\mu(z)$, then there are k_j eigenvalues converging almost surely to $H(\theta_j)$ iff $H'(\theta_j) > 0$.

Indeed, in that case, ν is semicircular, ω is invertible with inverse H.

Comments

Remark

This result illustrates that the free probabilistic interpretation of outliers, discovered in (CDFF 2010) generalizing the one in (BGN 2009), is a general principle.

Comments

Remark

This result illustrates that the free probabilistic interpretation of outliers, discovered in (CDFF 2010) generalizing the one in (BGN 2009), is a general principle.

Remark

It is noteworthy that, in this situation, a simple spike may create several outliers.

Sketch of proof-1

We use the following decomposition:

$$A_N = A'_N + A''_N,$$

$$A'_N = \text{Diag}(\alpha, \ldots, \alpha, \alpha_1^{(N)}, \ldots, \alpha_{N-r}^{(N)}),$$
$$A''_N = {}^tP\Theta P,$$

where P is the $r \times N$ matrix defined by

$$P = (I_r | 0_{r\times(N-r)}),$$

Θ is the $r \times r$ matrix

$$\Theta = \text{Diag}(\underbrace{\theta_1 - \alpha, \ldots, \theta_1 - \alpha}_{k_1}, \ldots, \underbrace{\theta_J - \alpha, \ldots, \theta_J - \alpha}_{k_J}),$$

and $\alpha \in \mathrm{supp}(\mu)$.

Sketch of proof-2

$$\det(\lambda I_N - X_N) = \det(\lambda I_N - (A'_N + U_N^* B_N U_N)) \det(I_N - R_N(\lambda)^t P \Theta P),$$

where

$$R_N(\lambda) = (\lambda I_N - (A'_N + U_N^* B_N U_N))^{-1}. \tag{2}$$

Using that, for rectangular matrices $X \in M_{N,r}(\mathbb{C})$, $Y \in M_{r,N}(\mathbb{C})$, one has $\det(I_N - XY) = \det(I_r - YX)$, one obtains:

$$\det(\lambda I_N - X_N)) = \det(\lambda I_N - (A'_N + U_N^* B_N U_N)) \det(I_r - P R_N(\lambda)^t P \Theta).$$

Hence, the outliers of X_N are precisely the zeros of $\det(M_N)$ outside the support of $\mu \boxplus \nu$, where

$$M_N := I_r - P R_N{}^t P \Theta. \tag{3}$$

Sketch of proof-3

Key point

Using Hurwitz's theorem, the zeros of $\det(M_N)$ will cluster towards those of $\det(M)$, where M is the almost sure uniform limit of M_N.

Sketch of proof-3

Key point

Using Hurwitz's theorem, the zeros of $\det(M_N)$ will cluster towards those of $\det(M)$, where M is the almost sure uniform limit of M_N.

- By concentration arguments, $M_N - I_r - P\mathbb{E}(R_N)^t P\Theta$ tends to 0 as N goes to infinity.

Sketch of proof-3

Key point

Using Hurwitz's theorem, the zeros of $\det(M_N)$ will cluster towards those of $\det(M)$, where M is the almost sure uniform limit of M_N.

- By concentration arguments, $M_N - I_r - P\mathbb{E}(R_N)^t P\Theta$ tends to 0 as N goes to infinity.
- It is known that $\mathbb{E}(R_N)$ is diagonal (Kargin 2011). Actually, it is a polynomial in A'_N.
  In particular, $P\mathbb{E}(R_N)^t P$ is a scalar matrix.

Sketch of proof-4

Define ω_N so that:

$$P\mathbb{E}(R_N)^t P = \frac{1}{\omega_N - \alpha} I_r.$$

Sketch of proof-4

Define ω_N so that:

$$P\mathbb{E}(R_N)^t P = \frac{1}{\omega_N - \alpha} I_r.$$

Then $(\omega_N)_{N \in \mathbb{N}}$ is a normal sequence of analytic functions, whose limit points l shall satisfy the subordination equation:

$$G_{\mu \boxplus \nu}(z) = G_\mu(l(z)),$$

which has the subordination map ω as a unique solution.

Sketch of proof-5

So M_N almost surely uniformly converges to:

$$M := I_r - \frac{1}{\omega - \alpha}\Theta. \tag{4}$$

Sketch of proof-5

So M_N almost surely uniformly converges to:

$$M := I_r - \frac{1}{\omega - \alpha}\Theta. \tag{4}$$

And z such that $det(M(z)) = 0$ are precisely solutions of $\omega(z) = \theta_j$ for some j, concluding the proof.

Thank you for your attention!