From random matrices to long range dependence

Arijit Chakrabarty

Joint work with Rajat S. Hazra and Deepayan Sarkar

July 1, 2014

The problem

- $(X_{i,j} : i, j \in \mathbb{Z})$ is a stationary, mean zero, variance one Gaussian process.
- Stationarity means that for $k, l \in \mathbb{Z}$,

$$(X_{i+k,j+l} : i, j \in \mathbb{Z}) \stackrel{d}{=} (X_{i,j} : i, j \in \mathbb{Z}) .$$

The problem

- $(X_{i,j} : i, j \in \mathbb{Z})$ is a stationary, mean zero, variance one Gaussian process.
- Stationarity means that for $k, l \in \mathbb{Z}$,

$$(X_{i+k,j+l} : i, j \in \mathbb{Z}) \stackrel{d}{=} (X_{i,j} : i, j \in \mathbb{Z}) \, .$$

- For $N \geq 1$, define a $N \times N$ matrix

$$W_N(i,j) := X_{i,j} + X_{j,i}, \, 1 \leq i, j \leq N \, .$$

- Denote

$$\mu_N := \mathrm{ESD}(W_N/\sqrt{N}) \, .$$

The problem

- $(X_{i,j} : i, j \in \mathbb{Z})$ is a stationary, mean zero, variance one Gaussian process.
- Stationarity means that for $k, l \in \mathbb{Z}$,

$$(X_{i+k,j+l} : i, j \in \mathbb{Z}) \stackrel{d}{=} (X_{i,j} : i, j \in \mathbb{Z}) \, .$$

- For $N \geq 1$, define a $N \times N$ matrix

$$W_N(i,j) := X_{i,j} + X_{j,i}, \; 1 \leq i, j \leq N \, .$$

- Denote

$$\mu_N := \mathrm{ESD}(W_N/\sqrt{N}) \, .$$

- **Goal:** To study the limit of μ_N as $N \to \infty$.

- Define

$$R(u,v) := \mathrm{E}\left(X_{0,0}X_{u,v}\right),\ u,v \in \mathbb{Z}\,.$$

- Herglotz theorem: There exists a finite measure ν on $(-\pi,\pi]^2$ satisfying

$$R(u,v) = \int_{(-\pi,\pi]^2} e^{\iota(ux+vy)}\nu(dx,dy),\ u,v \in \mathbb{Z}\,.$$

- Define

$$R(u,v) := \mathrm{E}\left(X_{0,0}X_{u,v}\right),\ u,v\in\mathbb{Z}\,.$$

- Herglotz theorem: There exists a finite measure ν on $(-\pi,\pi]^2$ satisfying

$$R(u,v) = \int_{(-\pi,\pi]^2} e^{\iota(ux+vy)}\nu(dx,dy),\ u,v\in\mathbb{Z}\,.$$

- Call ν the "spectral measure".

Assumption

- Write

$$\nu = \nu_{ac} + \nu_{cs} + \nu_d \,,$$

where

  - ν_{ac} is absolutely continuous,
  - ν_{cs} is continuous and singular,
  - ν_d is discrete.

Assumption

- Write
$$\nu = \nu_{ac} + \nu_{cs} + \nu_d \ ,$$
where
  - ν_{ac} is absolutely continuous,
  - ν_{cs} is continuous and singular,
  - ν_d is discrete.
- **Assumption**: $\nu_{cs} \equiv 0$.

Assumption

- Write
$$\nu = \nu_{ac} + \nu_{cs} + \nu_d \, ,$$
where
  - ν_{ac} is absolutely continuous,
  - ν_{cs} is continuous and singular,
  - ν_d is discrete.
- **Assumption**: $\nu_{cs} \equiv 0$.
- Let
$$\nu_{ac}(dx, dy) = f(x, y) dx\, dy \text{ on } (-\pi, \pi]^2 \, .$$

Results

Theorem

There exists a deterministic probability measure μ_f, determined solely by the spectral density f, such that

$$\mu_N \to \mu_f \,,$$

weakly in probability as $N \to \infty$.

Theorem

The second moment of the probability measure μ_f is given by

$$\int_{\mathbb{R}} x^2 \mu_f(dx) = 2 \int_{[-\pi,\pi]^2} f(x,y)\,dxdy\,.$$

Theorem

If

$$ess \inf f > 0\,,$$

then μ_f is absolutely continuous.

Theorem

1. For $m \geq 2$, the $(2m)$-th moment of μ_f is finite if $\|f\|_m < \infty$.
2. If $\|f\|_\infty < \infty$, then μ_f is compactly supported.

Stieltje's transform

Definition

Stieltje's transform of μ_f*:*

$$\mathcal{G}(z) := \int_{\mathbb{R}} \frac{1}{z-x} \mu_f(dx), z \in \mathbb{C} \setminus \mathbb{R} .$$

Stieltje's transform

Definition

Stieltje's transform of μ_f*:*

$$\mathcal{G}(z) := \int_{\mathbb{R}} \frac{1}{z-x} \mu_f(dx), z \in \mathbb{C} \setminus \mathbb{R}\,.$$

Theorem

Assume that $\|f\|_\infty < \infty$*. Then,*

$$\mathcal{G}(z) = \left[\int_{-\pi}^{\pi} \mathcal{H}(z,x)dx\right],\, z \in \mathbb{C}\,,$$

where $\mathcal{H}(z,x)$ *is the solution of the functional equation*

$$z\mathcal{H}(z,x) = 1 + \mathcal{H}(z,x) \int_{-\pi}^{\pi} \mathcal{H}(z,y) f(x,y) dy,\, z \in \mathbb{C} \setminus \mathbb{R}, |x| < \pi\,.$$

Theorem

If there exists a function r from $[-\pi,\pi]$ to $[0,\infty)$ such that

$$\frac{1}{2}\left[f(x,y)+f(y,x)\right]=r(x)r(y) \text{ for almost all } x,y\in[-\pi,\pi]\,,$$

then

$$\mu_f=\eta_r\boxtimes WSL(1)\,,$$

where η_r denotes the law of $2^{3/2}\pi r(U)$, U is an Uniform $(-\pi,\pi)$ random variable, and "$\boxtimes$" denotes the free multiplicative convolution.

Theorem

Assume that $(G_n : n \in \mathbb{Z})$ is a one-dimensional stationary Gaussian process with zero mean and positive variance, and whose spectral measure is absolutely continuous. Let $((G_{in} : n \in \mathbb{Z}) : i \in \mathbb{Z})$ be a family of i.i.d. copies of $(G_n : n \in \mathbb{Z})$. Define

$$X_{j,k} := G_{j-k,k},\, j,k \in \mathbb{Z}\,.$$

Then, $(X_{j,k} : j,k \in \mathbb{Z})$ is a stationary Gaussian process, and

$$\mu_f = WSL(2\mathrm{Var}(G_0))\,.$$

Effect of the discrete component

- A symmetric matrix A is to be thought of as a Hermitian operator $\overline{A}$ of finite rank acting on the first N coordinates of l^2.

Effect of the discrete component

- A symmetric matrix A is to be thought of as a Hermitian operator $\overline{A}$ of finite rank acting on the first N coordinates of l^2.
- If $\lambda_1 \leq \ldots \leq \lambda_N$ are the eigenvalues of A counted with multiplicity, then the spectrum of $\overline{A}$ is $\{0, \lambda_1, \ldots, \lambda_N\}$, where 0 has infinite multiplicity.

Eigen measure

- **Eigen measure** of A:

$$\mathrm{EM}(A) := \infty\delta_0 + \sum_{j=1}^{N} \delta_{\lambda_j}\,.$$

Eigen measure

- **Eigen measure** of A:

$$\mathrm{EM}(A) := \infty\delta_0 + \sum_{j=1}^{N} \delta_{\lambda_j} .$$

- The measure $\mathrm{EM}(A)$ is to be viewed as an element of the set $\mathcal{P}$ of point measures ξ of the form

$$\xi := \infty\delta_0 + \sum_{j=1}^{\infty} \delta_{\theta_j} ,$$

where $(\theta_j : j \geq 1)$ is some sequence of real numbers.

- $$\mathcal{C}_p := \left\{ \mu \in \mathcal{P} : \int_{\mathbb{R}} |x|^p \mu(dx) < \infty \right\}, \; p \in [1, \infty)\,.$$

-

$$\mathcal{C}_p := \left\{ \mu \in \mathcal{P} : \int_{\mathbb{R}} |x|^p \mu(dx) < \infty \right\}, \, p \in [1, \infty) \, .$$

- For $p \geq 1$ and $\xi \in \mathcal{C}_p$, there exist unique real numbers

$$\alpha_1(\xi) \geq \alpha_2(\xi) \geq \ldots \geq 0 \geq \ldots \alpha_{-2}(\xi) \geq \alpha_{-1}(\xi) \, ,$$

such that

$$\xi = \infty \delta_0 + \sum_{j \neq 0} \delta_{\alpha_j(\xi)} \, .$$

- $$\mathcal{C}_p := \left\{ \mu \in \mathcal{P} : \int_{\mathbb{R}} |x|^p \mu(dx) < \infty \right\}, \, p \in [1, \infty) \, .$$

- For $p \geq 1$ and $\xi \in \mathcal{C}_p$, there exist unique real numbers

$$\alpha_1(\xi) \geq \alpha_2(\xi) \geq \ldots \geq 0 \geq \ldots \alpha_{-2}(\xi) \geq \alpha_{-1}(\xi) \, ,$$

such that

$$\xi = \infty \delta_0 + \sum_{j \neq 0} \delta_{\alpha_j(\xi)} \, .$$

- $$d_p(\xi_1, \xi_2) := \left[\sum_{j \neq 0} |\alpha_j(\xi_1) - \alpha_j(\xi_2)|^p \right]^{1/p} , \, \xi_1, \xi_2 \in \mathcal{C}_p \, .$$

Results

Theorem

There exists a random point measure ξ which is almost surely in $\mathcal{C}_2$ such that

$$d_4\left(\mathrm{EM}(W_N/N), \xi\right) \xrightarrow{P} 0\,,$$

as $N \to \infty$. Furthermore, the distribution of ξ is determined by ν_d.

Results

Theorem

There exists a random point measure ξ which is almost surely in $\mathcal{C}_2$ such that

$$d_4\left(\mathrm{EM}(W_N/N), \xi\right) \xrightarrow{P} 0\,,$$

as $N \to \infty$. Furthermore, the distribution of ξ is determined by ν_d.

Remark

If $f \equiv 0$, then d_4 can be improved to d_2.

Theorem

If $\nu_d((-\pi,\pi]^2) > 0$, then the random variable

$$\int_{\mathbb{R}} x^2 \xi(dx)$$

is positive almost surely, and non-degenerate.

Long range dependence

1.	The component ν_{ac} determines the limiting ESD	The component ν_d determines the limiting EM

Long range dependence

1.	The component ν_{ac} determines the limiting ESD μ_f, of $W_N/\sqrt{N}$.	The component ν_d determines the limiting EM ξ, of W_N/N .

Long range dependence

1.	The component ν_{ac} determines the limiting ESD μ_f, of $W_N/\sqrt{N}$.	The component ν_d determines the limiting EM ξ, of W_N/N .
2.	The measure μ_f is deterministic.	The measure ξ is random.

Long range dependence

1.	The component ν_{ac} determines the limiting ESD μ_f, of $W_N/\sqrt{N}$.	The component ν_d determines the limiting EM ξ, of W_N/N .
2.	The measure μ_f is deterministic.	The measure ξ is random.

Definition

A mean zero stationary Gaussian process with positive variance indexed by $\mathbb{Z}^2$ is **short range dependent** *if the corresponding spectral measure is absolutely continuous, and the same is* **long range dependent** *if the spectral measure is discrete, that is, supported on a countable set.*

Example 1

- $(G_n : n \in \mathbb{Z})$: a zero mean stationary Gaussian process with spectral density $|x|^{-1/2}$,
- $((G_{in} : n \in \mathbb{Z}) : i \in \mathbb{Z})$: a family of i.i.d. copies of $(G_n : n \in \mathbb{Z})$,
- $$X_{j,k} := G_{j-k,k},\, j,k \in \mathbb{Z}\,.$$

Example 1

- $(G_n : n \in \mathbb{Z})$: a zero mean stationary Gaussian process with spectral density $|x|^{-1/2}$,
- $((G_{in} : n \in \mathbb{Z}) : i \in \mathbb{Z})$: a family of i.i.d. copies of $(G_n : n \in \mathbb{Z})$,
- $$X_{j,k} := G_{j-k,k},\, j, k \in \mathbb{Z}\,.$$
- $$\mu_f = WSL(2\mathrm{Var}(G_0))\,.$$

Example 2

- $$f(x,y) := \mathbf{1}(-\pi/2 \le x, y \le \pi/2),\ -\pi \le x, y \le \pi\,.$$

Example 2

- $$f(x, y) := \mathbf{1}(-\pi/2 \le x, y \le \pi/2),\ -\pi \le x, y \le \pi\,.$$

- $$\mu_f = \eta_r \boxtimes WSL(1)\,,$$

  where
  - η_r is the law of $2^{3/2}\pi\mathbf{1}(|U| \le \pi/2)$,
  - $U \sim (-\pi, \pi)$.

Example 2

- $$f(x, y) := \mathbf{1}(-\pi/2 \le x, y \le \pi/2),\ -\pi \le x, y \le \pi\,.$$

- $$\mu_f = \eta_r \boxtimes WSL(1)\,,$$

  where
  - η_r is the law of $2^{3/2}\pi\mathbf{1}(|U| \le \pi/2)$,
  - $U \sim (-\pi, \pi)$.
- μ_f is the law of $2\pi BW$ where
  - $B \sim$Bernoulli $(1/2)$,
  - $W \sim WSL(1)$,
  - B, W are **classically** independent.

Example 3

- $$f(x,y) := |xy|^{-1/2},\ -\pi \le x, y \le \pi\,.$$
- $$\mu_f = \eta_r \boxtimes WSL(1)\,.$$
- $$\int_{\mathbb{R}} x^4 \mu_f(dx) = \infty\,.$$

The main ingredient

Fact

Let $(X_{j,k} : j, k \in \mathbb{Z})$ *be a stationary mean zero Gaussian process. Then, there exist* $c_{j,k} \in \mathbb{R}$ *satisfying*

$$\sum_{j,k} c_{j,k}^2 < \infty \,,$$

and

$$(X_{j,k} : j, k \in \mathbb{Z}) \stackrel{d}{=} \left(\sum_{l,m} c_{l,m} G_{j-l,k-m} : j, k \in \mathbb{Z} \right) \,,$$

if and only if, the spectral measure of $(X_{j,k})$ *is absolutely continuous.*

Proof of the main result

Theorem

There exists a deterministic probability measure μ_f, determined solely by the spectral density f, such that

$$\mu_N \to \mu_f \,,$$

weakly in probability as $N \to \infty$, where

$$\mu_N := \mathrm{ESD}(W_N/\sqrt{N})\,.$$

Proof

- Assume that $\nu \equiv \nu_{ac}$.

Proof

- Assume that $\nu \equiv \nu_{ac}$.
-

$$(X_{j,k} : j,k \in \mathbb{Z}) \stackrel{d}{=} \left(\sum_{l,m} c_{l,m} G_{j-l,k-m} : j,k \in \mathbb{Z} \right) ,$$

Proof

- Assume that $\nu \equiv \nu_{ac}$.
- $$(X_{j,k} : j, k \in \mathbb{Z}) \stackrel{d}{=} \left(\sum_{l,m} c_{l,m} G_{j-l,k-m} : j, k \in \mathbb{Z} \right),$$
- Usual moments method works for
$$\sum_{l,m=-n}^{n} c_{l,m} G_{j-l,k-m}$$
for fixed n.

Proof

- Assume that $\nu \equiv \nu_{ac}$.
- $$(X_{j,k} : j, k \in \mathbb{Z}) \stackrel{d}{=} \left(\sum_{l,m} c_{l,m} G_{j-l,k-m} : j, k \in \mathbb{Z} \right) ,$$
- Usual moments method works for
$$\sum_{l,m=-n}^{n} c_{l,m} G_{j-l,k-m}$$
for fixed n.
- Hoffman-Wielandt inequality completes the proof.

Free probability

Theorem

Let μ be a probability measure on $\mathbb{R}$ such that

$$\mathit{Support}(\mu) \subset [\delta, \infty) \text{ for some } \delta > 0\,,$$

and

$$\int_0^\infty x\mu(dx) < \infty\,.$$

Then, there exists a probability measure ν on $\mathbb{R}$ and $\varepsilon > 0$ such that

$$WSL(1) \boxtimes \mu = WSL(\varepsilon) \boxplus \nu\,.$$

Corollary

If μ is as in the previous result, then WSL $\boxtimes \mu$ is absolutely continuous.

Future work

- The ν_{cs} component.

Future work

- The ν_{cs} component.
- Growth rate of the largest eigenvalue, especially when the LSD has unbounded support.

Future work

- The ν_{cs} component.
- Growth rate of the largest eigenvalue, especially when the LSD has unbounded support.
- The asymmetric case.