On the empirical spectral distribution for sample covariance matrices with long memory

Florence Merlevède
Université Paris-Est-Marne-La-Vallée (UPEM)

Joint work with M. Peligrad

Random matrices and their applications
Hong-Kong, January 2015

Introduction

- Consider N independent copies $(X_{ij})_{j\in\mathbb{Z}}$, $i = 1, \dots, N$ of a stationary sequence $(X_i)_{i\in\mathbb{Z}}$ of real-valued r.v's and consider the $N \times p$ matrix

$$\mathcal{X}_{N,p} = (X_{ij})_{1\le i\le N, 1\le j\le p}$$

- Consider N independent copies $(X_{ij})_{j\in\mathbb{Z}}$, $i=1,\dots,N$ of a stationary sequence $(X_i)_{i\in\mathbb{Z}}$ of real-valued r.v's and consider the $N\times p$ matrix

$$\mathcal{X}_{N,p}=(X_{ij})_{1\leq i\leq N,1\leq j\leq p}$$

- Define now the symmetric matrix $\mathbf{B}_N$ of order p by

$$\mathbf{B}_N=\frac{1}{N}\mathcal{X}_{N,p}^T\mathcal{X}_{N,p}:=\frac{1}{N}\sum_{i=1}^{N}C_i^TC_i$$

  where $C_i=(X_{i1},\dots,X_{ip})$. $\mathbf{B}_N$ is usually called the sample covariance matrix (or Gram matrix) associated with the process (X_{ij}).

- Question: What can we say about the spectrum of B_N when $\lim\frac{p}{N}\to c\in(0,\infty)$?

Introduction

- Consider N independent copies $(X_{ij})_{j\in\mathbb{Z}}$, $i=1,\ldots,N$ of a stationary sequence $(X_i)_{i\in\mathbb{Z}}$ of real-valued r.v's and consider the $N\times p$ matrix

$$\mathcal{X}_{N,p}=(X_{ij})_{1\le i\le N,1\le j\le p}$$

- Define now the symmetric matrix $\mathbf{B}_N$ of order p by

$$\mathbf{B}_N=\frac{1}{N}\mathcal{X}_{N,p}^T\mathcal{X}_{N,p}:=\frac{1}{N}\sum_{i=1}^{N}C_i^TC_i$$

  where $C_i=(X_{i1},\ldots,X_{ip})$. $\mathbf{B}_N$ is usually called the sample covariance matrix (or Gram matrix) associated with the process (X_{ij}).

- Question: What can we say about the spectrum of B_N when $\lim\frac{p}{N}\to c\in(0,\infty)$?
- Let us look at the spectral measure of $\mathbf{B}_N$

$$\mu_{\mathbf{B}_N}:=\frac{1}{p}\sum_{i=1}^{p}\delta_{\lambda_i}\text{ where }\lambda_i=\lambda_i(\mathbf{B}_N)$$

Marčenko-Pastur (1967)

Marčenko-Pastur (1967)

- Assume that $(X_i)_{i\in\mathbb{Z}}$ is a sequence of iid, mean zero and with variance 1 r.v.'s (so the entries of $\mathcal{X}_{N,p}$ are iid). Assume that $\lim \frac{p}{N} \to c \in (0,\infty)$. Then, with probability 1, $\mu_{\mathbf{B}_N}$ converge in law to a non random probability measure μ_{MP} whose density is given by

$$f(x) = \frac{1}{2\pi cx}\sqrt{(b-x)(x-a)}\mathbf{1}_{[a,b]}(x) + \mathbf{1}_{]1,\infty)}(c)(1-c^{-1})\mathbf{1}_{x=0}$$

with $a = (1-\sqrt{c})^2$, $b = (1+\sqrt{c})^2$.

Marčenko-Pastur (1967)

- Assume that $(X_i)_{i\in\mathbb{Z}}$ is a sequence of iid, mean zero and with variance 1 r.v.'s (so the entries of $\mathcal{X}_{N,p}$ are iid). Assume that $\lim \frac{p}{N} \to c \in (0,\infty)$. Then, with probability 1, $\mu_{\mathbf{B}_N}$ converge in law to a non random probability measure μ_{MP} whose density is given by

$$f(x) = \frac{1}{2\pi c x}\sqrt{(b-x)(x-a)}\mathbf{1}_{[a,b]}(x) + \mathbf{1}_{]1,\infty)}(c)(1-c^{-1})\mathbf{1}_{x=0}$$

  with $a=(1-\sqrt{c})^2$, $b=(1+\sqrt{c})^2$.

- Since 1967, there has been a great amount of work to relax the assumption on the independence of the entries of $\mathcal{X}_{N,p}$, and in particular the independence structure in the rows of $\mathcal{X}_{N,p}$.

A first extension of the Marčenko-Pastur's result

- Yin (1986) and Silverstein (1995). They consider

$$\Sigma_N = \frac{1}{\sqrt{N}} R_p^{1/2} \mathcal{X}_{N,p}^T$$

where $\mathcal{X}_{N,p} := (X_{ij})_{1 \leq i \leq N, 1 \leq j \leq p}$, the X_{ij}'s are iid with variance 1, and R_p is a Hermitian non negative definite random matrix of size p independent of $\mathcal{X}_{N,p}$. Let $\mathbf{B}_N = \Sigma_N \Sigma_N^T$.

A first extension of the Marčenko-Pastur's result

- Yin (1986) and Silverstein (1995). They consider

$$\Sigma_N = \frac{1}{\sqrt{N}} R_p^{1/2} \mathcal{X}_{N,p}^T$$

where $\mathcal{X}_{N,p} := (X_{ij})_{1 \le i \le N, 1 \le j \le p}$, the X_{ij}'s are iid with variance 1, and R_p is a Hermitian non negative definite random matrix of size p independent of $\mathcal{X}_{N,p}$. Let $\mathbf{B}_N = \Sigma_N \Sigma_N^T$.

- Silverstein (1995): If $\lim_{n\to\infty} p/N = c \in (0, \infty)$ and if with probability one, μ_{R_p} converges in law to a non random probability measure μ_H then, almost surely, $\mu_{\mathbf{B}_n}$ converges in law to a non random probability measure characterized via its Stieltjes transform. More precisely, for any $z \in \mathbb{C}^+$

$$S_{\mu_{\mathbf{B}_N}}(z) = \int \frac{1}{x - z} d\mu_{\mathbf{B}_N}(x) = \frac{1}{p} \mathrm{Tr}(\mathbf{B}_N - z\mathbf{I}_p)^{-1} \to S(z) \text{ almost surely}$$

where

$$S(z) = \int \frac{d\mu_H(x)}{-z + (1-c)x - zcxS(z)}$$

A second extension of the Marčenko-Pastur's result: Bai-Zhou (2008)

For $1 \leq i \leq N$, let $C_i = (X_{i1}, \ldots, X_{ip})$ and $\mathbf{B}_N = \frac{1}{N}\sum_{i=1}^N C_i^T C_i$. Assume that the C_i's are independent, and that

- For all i, $\mathbb{E}(X_{ik}X_{i\ell}) = \gamma_{k,\ell}$ and for any deterministic matrix $p \times p$, $R = (r_{k\ell})$, with bounded spectral norm

$$\mathbb{E}\left|C_i R C_i^T - \mathrm{Tr}(R\Gamma_p)\right|^2 = o(N^2) \quad \text{where } \Gamma_p = (\gamma_{k,\ell})$$

- $\lim_{N\to\infty} p/N = c \in (0, \infty)$
- The spectral norm of Γ_p is uniformly bounded and μ_{Γ_p} converges in law to μ_H.

Then, almost surely, $\mu_{\mathbf{B}_N}$ converges in law to a non random probability measure whose Stieltjes transform $S = S(z)$ satisfies the equation : for all $z \in \mathbb{C}^+$

$$z = -\frac{1}{\underline{S}} + c\int \frac{t}{1+\underline{S}t} d\mu_H(t)\,,$$

where $\underline{S}(z) := -(1-c)/z + cS(z)$.

Applications

- The assumption

$$\mathbb{E}\left|C_i R C_i^T - \mathrm{Tr}(R\Gamma_p)\right|^2 = o(N^2) \quad \text{where } \Gamma_p = (\gamma_{k,\ell})$$

for any $p \times p$ deterministic matrix, $R = (r_{k,\ell})$, with bounded spectral norm, is verified as soon as

  - $N^{-1} \max_{k \neq \ell} \mathbb{E}\big(X_{ik}X_{i\ell} - \gamma_{k,\ell}\big)^2 \to 0$ uniformly in $i \leq N$
  - $N^{-2} \sum_{\Lambda} \Big(\mathbb{E}(X_{ik}X_{i\ell} - \gamma_{k,\ell})(X_{ik'}X_{i\ell'} - \gamma_{k',\ell'})\Big)^2 \to 0$ uniformly in $i \leq N$

where $\Lambda = \{(k,\ell,k',\ell') : 1 \leq k,\ell,k',\ell' \leq p\} \backslash \{(k,\ell,k',\ell') : k = k' \neq \ell = \ell' \text{ or } k = \ell' \neq k' = \ell\}$.

Applications

- The assumption

$$\mathbb{E}\left|C_i R C_i^T - \mathrm{Tr}(R\Gamma_p)\right|^2 = o(N^2) \quad \text{where } \Gamma_p = (\gamma_{k,\ell})$$

for any $p \times p$ deterministic matrix, $R = (r_{k,\ell})$, with bounded spectral norm, is verified as soon as

  - $N^{-1} \max_{k \neq \ell} \mathbb{E}\big(X_{ik}X_{i\ell} - \gamma_{k,\ell}\big)^2 \to 0$ uniformly in $i \leq N$
  - $N^{-2} \sum_{\Lambda} \Big(\mathbb{E}(X_{ik}X_{i\ell} - \gamma_{k,\ell})(X_{ik'}X_{i\ell'} - \gamma_{k',\ell'})\Big)^2 \to 0$ uniformly in $i \leq N$

  where $\Lambda = \{(k,\ell,k',\ell') : 1 \leq k,\ell,k',\ell' \leq p\} \backslash \{(k,\ell,k',\ell') : k = k' \neq \ell = \ell' \text{ or } k = \ell' \neq k' = \ell\}$.

- In their paper, Bai and Zhou applied their result to get the limiting spectral distributions of Spearman's rank correlation matrices, of sample correlation matrices from finite population and of sample covariance matrices generated by causal AR(1) models.

Application to sample covariance matrices generated by causal linear processes

- Yao ('12) and Pan *et al.* ('14): Let $\mathbf{B}_N = \frac{1}{N}\sum_{i=1}^N C_i^T C_i$ where the $C_i = (X_{i1}, \ldots, X_{ip})$'s are independent copies of $C = (X_1, \ldots, X_p)$ with $X_k = \sum_{j\geq 0} a_j \varepsilon_{k-j}$ where the ε_k's are iid, centered and in $\mathbf{L}^2$, and $\sum_{k\geq 1}|a_k| < \infty$. Then, when $\lim_{N\to\infty} p/N = c \in (0,\infty)$, with probability one, $\mu_{\mathbf{B}_N}$ converges in law to a non random probability measure whose Stieltjes transform $S = S(z)$ satisfies the equation

$$z = -\frac{1}{\underline{S}} + \frac{c}{2\pi}\int_{-\pi}^{\pi} \frac{1}{\underline{S} + (2\pi f(\lambda))^{-1}} d\lambda\,,$$

where $f(\cdot)$ is the spectral density of $(X_k)_{k\in\mathbb{Z}}$. The limiting spectral distribution has compact support.

Application to sample covariance matrices generated by causal linear processes

- Yao ('12) and Pan *et al.* ('14): Let $\mathbf{B}_N = \frac{1}{N}\sum_{i=1}^{N} C_i^T C_i$ where the $C_i = (X_{i1}, \dots, X_{ip})$'s are independent copies of $C = (X_1, \dots, X_p)$ with $X_k = \sum_{j\geq 0} a_j \varepsilon_{k-j}$ where the ε_k's are iid, centered and in $\mathbf{L}^2$, and $\sum_{k\geq 1} |a_k| < \infty$. Then, when $\lim_{N\to\infty} p/N = c \in (0,\infty)$, with probability one, $\mu_{\mathbf{B}_N}$ converges in law to a non random probability measure whose Stieltjes transform $S = S(z)$ satisfies the equation

$$z = -\frac{1}{\underline{S}} + \frac{c}{2\pi}\int_{-\pi}^{\pi} \frac{1}{\underline{S} + (2\pi f(\lambda))^{-1}} d\lambda \,,$$

  where $f(\cdot)$ is the spectral density of $(X_k)_{k\in\mathbb{Z}}$. The limiting spectral distribution has compact support.

- The assumption $\mathbb{E}\big|C_i R C_i^T - \mathrm{Tr}(R\Gamma_p)\big| = o(N^2)$ can be hard to verify for non linear time series (ex of ARCH models) or requires rate of convergence of mixing coefficients.

Application to sample covariance matrices generated by causal linear processes

- Yao ('12) and Pan *et al.* ('14): Let $\mathbf{B}_N = \frac{1}{N}\sum_{i=1}^{N} C_i^T C_i$ where the $C_i = (X_{i1}, \dots, X_{ip})$'s are independent copies of $C = (X_1, \dots, X_p)$ with $X_k = \sum_{j\geq 0} a_j \varepsilon_{k-j}$ where the ε_k's are iid, centered and in $\mathbf{L}^2$, and $\sum_{k\geq 1} |a_k| < \infty$. Then, when $\lim_{N\to\infty} p/N = c \in (0,\infty)$, with probability one, $\mu_{\mathbf{B}_N}$ converges in law to a non random probability measure whose Stieltjes transform $S = S(z)$ satisfies the equation

$$z = -\frac{1}{\underline{S}} + \frac{c}{2\pi}\int_{-\pi}^{\pi} \frac{1}{\underline{S} + (2\pi f(\lambda))^{-1}} d\lambda \,,$$

  where $f(\cdot)$ is the spectral density of $(X_k)_{k\in\mathbb{Z}}$. The limiting spectral distribution has compact support.

- The assumption $\mathbb{E}\big|C_i R C_i^T - \mathrm{Tr}(R\Gamma_p)\big| = o(N^2)$ can be hard to verify for non linear time series (ex of ARCH models) or requires rate of convergence of mixing coefficients.

- Does the spectral distribution of sample covariance matrices generated by causal linear processes with innovations in $\mathbb{L}^2$ and $\sum_{k\geq 1} a_k^2 < \infty$ admits a limit?

Main result: M. & Peligrad (2014)

- Consider N independent copies $(X_{ij})_{j\in\mathbb{Z}}$, $i = 1, \dots, N$ of a stationary sequence $(X_i)_{i\in\mathbb{Z}}$ of real-valued random variables centered and in $\mathbb{L}^2$ and that satisfies the following regularity conditions:

$$\mathbb{E}(X_0|\mathcal{G}_{-\infty}) = 0 \text{ a.s.}$$

and for every integer k

$$\mathbb{E}(X_0 X_k|\mathcal{G}_{-\infty}) = \mathbb{E}(X_0 X_k) \text{ a.s.}$$

where $\mathcal{G}_{-\infty} = \bigcap_{k\in\mathbb{Z}} \mathcal{G}_k$ with $\mathcal{G}_k = \sigma(X_j, j \le k)$.

Main result: M. & Peligrad (2014)

- Consider N independent copies $(X_{ij})_{j\in\mathbb{Z}}$, $i=1,\dots,N$ of a stationary sequence $(X_i)_{i\in\mathbb{Z}}$ of real-valued random variables centered and in $\mathbb{L}^2$ and that satisfies the following regularity conditions:

$$\mathbb{E}(X_0|\mathcal{G}_{-\infty}) = 0 \text{ a.s.}$$

  and for every integer k

$$\mathbb{E}(X_0X_k|\mathcal{G}_{-\infty}) = \mathbb{E}(X_0X_k) \text{ a.s.}$$

  where $\mathcal{G}_{-\infty} = \bigcap_{k\in\mathbb{Z}}\mathcal{G}_k$ with $\mathcal{G}_k = \sigma(X_j, j\le k)$.

- Assume $p/N \to c \in (0,\infty)$. Then, almost surely, $\mu_{\mathbf{B}_N}$ converges in law to a non random probability measure whose Stieltjes transform $S=S(z)$ satisfies the equation : for all $z\in\mathbb{C}^+$

$$z = -\frac{1}{\underline{S}} + \frac{c}{2\pi}\int_{-\pi}^{\pi}\frac{1}{\underline{S}+(2\pi f(\lambda))^{-1}}\,d\lambda\,,$$

  where $\underline{S} := -(1-c)/z + cS$ and $f(\cdot)$ is the spectral density of $(X_k)_{k\in\mathbb{Z}}$.

On the regularity conditions

Recall the regularity conditions:

$$\mathbb{E}(X_0|\mathcal{G}_{-\infty}) = 0 \text{ a.s.} \tag{1}$$

and for every nonnegative integer k

$$\mathbb{E}(X_0 X_k|\mathcal{G}_{-\infty}) = \mathbb{E}(X_0 X_k) \text{ a.s.} \tag{2}$$

- (1) implies that the process $(X_k)_{k\in\mathbb{Z}}$ is purely non deterministic. Therefore, by a result of Szegö, the spectral density of $(X_k)_{k\in\mathbb{Z}}$ exists and if X_0 is non degenerate, f cannot vanish on a set of positive measure.

On the regularity conditions

Recall the regularity conditions:

$$\mathbb{E}(X_0|\mathcal{G}_{-\infty}) = 0 \text{ a.s.} \tag{1}$$

and for every nonnegative integer k

$$\mathbb{E}(X_0 X_k|\mathcal{G}_{-\infty}) = \mathbb{E}(X_0 X_k) \text{ a.s.} \tag{2}$$

- (1) implies that the process $(X_k)_{k\in\mathbb{Z}}$ is purely non deterministic. Therefore, by a result of Szegö, the spectral density of $(X_k)_{k\in\mathbb{Z}}$ exists and if X_0 is non degenerate, f cannot vanish on a set of positive measure.
- If the left tail sigma field $\mathcal{G}_{-\infty}$ is trivial then (1) and (2) hold. This is the case when (X_k) is a function of an iid r.v.'s sequence, so in this situation no condition is required except the process to be in $\mathbb{L}^2$.

On the regularity conditions

Recall the regularity conditions:

$$\mathbb{E}(X_0|\mathcal{G}_{-\infty}) = 0 \text{ a.s.} \tag{1}$$

and for every nonnegative integer k

$$\mathbb{E}(X_0 X_k|\mathcal{G}_{-\infty}) = \mathbb{E}(X_0 X_k) \text{ a.s.} \tag{2}$$

- (1) implies that the process $(X_k)_{k\in\mathbb{Z}}$ is purely non deterministic. Therefore, by a result of Szegö, the spectral density of $(X_k)_{k\in\mathbb{Z}}$ exists and if X_0 is non degenerate, f cannot vanish on a set of positive measure.
- If the left tail sigma field $\mathcal{G}_{-\infty}$ is trivial then (1) and (2) hold. This is the case when (X_k) is a function of an iid r.v.'s sequence, so in this situation no condition is required except the process to be in $\mathbb{L}^2$.
- It follows that if $X_k = \sum_{j\geq 0} a_j \varepsilon_{k-j}$ where the ε_k's are iid, centered and in $\mathbf{L}^2$, and $\sum_{k\geq 0} a_k^2 < \infty$, the result applies.

Other situations where the left tail sigma field is trivial

- Let $\mathcal{G}^n = \sigma(X_k, k \geq n)$ and define

$$\alpha_n = \alpha(\mathcal{G}_0, \mathcal{G}^n) \ \text{ and } \rho_n = \rho(\mathcal{G}_0, \mathcal{G}^n) \,.$$

where for two two sigma algebras $\mathcal{A}$ and $\mathcal{B}$,

$$\alpha(\mathcal{A}, \mathcal{B}) = \sup\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \mathcal{A},\ B \in \mathcal{B}\}$$

and

$$\rho(\mathcal{A}, \mathcal{B}) = \sup\{\mathrm{Cov}(X, Y)/\|X\|_2\|Y\|_2 : X \in \mathbb{L}^2(\mathcal{A}),\ Y \in \mathbb{L}^2(\mathcal{B})\}.$$

Other situations where the left tail sigma field is trivial

- Let $\mathcal{G}^n = \sigma(X_k, k \geq n)$ and define
$$\alpha_n = \alpha(\mathcal{G}_0, \mathcal{G}^n) \text{ and } \rho_n = \rho(\mathcal{G}_0, \mathcal{G}^n) .$$
where for two two sigma algebras $\mathcal{A}$ and $\mathcal{B}$,
$$\alpha(\mathcal{A}, \mathcal{B}) = \sup\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \mathcal{A}, B \in \mathcal{B}\}$$
and
$$\rho(\mathcal{A}, \mathcal{B}) = \sup\{\mathrm{Cov}(X, Y)/\|X\|_2\|Y\|_2 : X \in \mathbb{L}^2(\mathcal{A}), Y \in \mathbb{L}^2(\mathcal{B})\}.$$
- If $\alpha_n \to 0$ or $\lim_{n\to\infty} \rho_n < 1$, then the left tail sigma field $\mathcal{G}_{-\infty}$ is trivial (see Bradley (2005) and (2007)).

Other situations where the left tail sigma field is trivial

- Let $\mathcal{G}^n = \sigma(X_k, k \geq n)$ and define
$$\alpha_n = \alpha(\mathcal{G}_0, \mathcal{G}^n) \;\text{ and } \rho_n = \rho(\mathcal{G}_0, \mathcal{G}^n)\,.$$
where for two two sigma algebras $\mathcal{A}$ and $\mathcal{B}$,
$$\alpha(\mathcal{A}, \mathcal{B}) = \sup\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)| : A \in \mathcal{A},\ B \in \mathcal{B}\}$$
and
$$\rho(\mathcal{A}, \mathcal{B}) = \sup\{\mathrm{Cov}(X, Y)/\|X\|_2\|Y\|_2 : X \in \mathbb{L}^2(\mathcal{A}),\ Y \in \mathbb{L}^2(\mathcal{B})\}.$$
- If $\alpha_n \to 0$ or $\lim_{n\to\infty} \rho_n < 1$, then the left tail sigma field $\mathcal{G}_{-\infty}$ is trivial (see Bradley (2005) and (2007)).
- Triviality of $\mathcal{G}_{-\infty}$ is not necessary for the regularity conditions to hold. For instance, it is enough that
$$\alpha_{2,n} = \sup_{k\geq 0} \alpha(\mathcal{G}_{-n}, \sigma(X_0, X_k)) \to 0$$
and the variables to be centered and in $\mathbb{L}^2$, for the regularity conditions to hold

Sketch of proof

The result will follow if we can prove that: for any $z \in \mathbb{C}^+$,

Sketch of proof

The result will follow if we can prove that: for any $z \in \mathbb{C}^+$,

- $S_{\mu_{\mathbf{B}_N}}(z) - \mathbf{E} S_{\mu_{\mathbf{B}_N}}(z) \to 0$ a.s.

Sketch of proof

The result will follow if we can prove that: for any $z \in \mathbb{C}^+$,

- $S_{\mu_{\mathbf{B}_N}}(z) - \mathbf{E} S_{\mu_{\mathbf{B}_N}}(z) \to 0$ a.s.
- $\mathbf{E} S_{\mu_{\mathbf{B}_N}}(z) - \mathbf{E} S_{\mu_{\mathbf{G}_N}}(z) \to 0$ where

$$\mathbf{G}_N = \frac{1}{N} \sum_{i=1}^{N} G_i^T G_i = \frac{1}{N} \mathcal{G}_{N,p}^T \mathcal{G}_{N,p} \,,$$

with $G_i = (Y_{i1}, \ldots, Y_{ip})$ where $(Y_{ik})_{k \in \mathbb{Z}}$, $i = 1, \cdots, N$ are N independent copies of a centered real-valued Gaussian process $(Y_k)_{k \in \mathbb{Z}}$ such that for all integers i, j

$$\mathrm{Cov}(Y_i, Y_j) = \mathrm{Cov}(X_i, X_j)$$

We can take (Y_{ik}) independent of (X_{ik}).

The result will follow if we can prove that: for any $z \in \mathbb{C}^+$,

- $S_{\mu_{\mathbf{B}_N}}(z) - \mathbf{E} S_{\mu_{\mathbf{B}_N}}(z) \to 0$ a.s.
- $\mathbf{E} S_{\mu_{\mathbf{B}_N}}(z) - \mathbf{E} S_{\mu_{\mathbf{G}_N}}(z) \to 0$ where

$$\mathbf{G}_N = \frac{1}{N} \sum_{i=1}^{N} G_i^T G_i = \frac{1}{N} \mathcal{G}_{N,p}^T \mathcal{G}_{N,p} \,,$$

with $G_i = (Y_{i1}, \ldots, Y_{ip})$ where $(Y_{ik})_{k \in \mathbb{Z}}$, $i = 1, \cdots, N$ are N independent copies of a centered real-valued Gaussian process $(Y_k)_{k \in \mathbb{Z}}$ such that for all integers i, j

$$\mathrm{Cov}(Y_i, Y_j) = \mathrm{Cov}(X_i, X_j)$$

We can take (Y_{ik}) independent of (X_{ik}).

- $\mathbf{E} S_{\mu_{\mathbf{G}_N}}(z) \to S(z)$.

- The fact that $S_{\mu_{\mathbf{B}_N}}(z) - \mathbf{E}S_{\mu_{\mathbf{B}_N}}(z) \to 0$ a.s. comes directly from concentration results for spectral measures (see for instance Guntuboyina-Leeb (2009)).

- The fact that $S_{\mu_{\mathbf{B}_N}}(z) - \mathbf{E}S_{\mu_{\mathbf{B}_N}}(z) \to 0$ a.s. comes directly from concentration results for spectral measures (see for instance Guntuboyina-Leeb (2009)).
- To prove that $\mathbf{E}S_{\mu_{\mathbf{B}_N}}(z) - \mathbf{E}S_{\mu_{\mathbf{G}_N}}(z) \to 0$, notice that it is enough to show that

$$\mathbf{E}S_{\mu_{\mathbb{X}_n}}(z) - \mathbf{E}S_{\mu_{\mathbb{Y}_n}}(z) \to 0 \quad (*)$$

where $n = N + p$,

$$\mathbb{X}_n = \frac{1}{\sqrt{N}} \begin{pmatrix} \mathbf{0}_{p,p} & \mathcal{X}_{N,p}^T \\ \mathcal{X}_{N,p} & \mathbf{0}_{N,N} \end{pmatrix} \text{ and } \mathbb{Y}_n = \frac{1}{\sqrt{N}} \begin{pmatrix} \mathbf{0}_{p,p} & \mathcal{G}_{N,p}^T \\ \mathcal{G}_{N,p} & \mathbf{0}_{N,N} \end{pmatrix}.$$

Indeed

$$S_{\mu_{\mathbf{B}_N}}(z) = z^{-1/2} \frac{n}{2p} S_{\mu_{\mathbb{X}_n}}(z^{1/2}) + \frac{N-p}{2pz}$$

Note that $\mathbb{X}_n := N^{-1/2}[x_{ij}^{(n)}]_{i,j=1}^n$ and $\mathbb{Y}_n := N^{-1/2}[y_{ij}^{(n)}]_{i,j=1}^n$ where

$$z_{ij}^{(n)} = \begin{cases} Z_{i-p,j} \mathbf{1}_{i \geq p+1} \mathbf{1}_{1 \leq j \leq p} & \text{if } 1 \leq j \leq i \leq n \\ z_{ji}^{(n)} & \text{if } 1 \leq i < j \leq n \end{cases}.$$

- The fact that $S_{\mu_{\mathbf{B}_N}}(z) - \mathbf{E}S_{\mu_{\mathbf{B}_N}}(z) \to 0$ a.s. comes directly from concentration results for spectral measures (see for instance Guntuboyina-Leeb (2009)).
- To prove that $\mathbf{E}S_{\mu_{\mathbf{B}_N}}(z) - \mathbf{E}S_{\mu_{\mathbf{G}_N}}(z) \to 0$, notice that it is enough to show that

$$\mathbf{E}S_{\mu_{\mathbb{X}_n}}(z) - \mathbf{E}S_{\mu_{\mathbb{Y}_n}}(z) \to 0 \quad (*)$$

where $n = N + p$,

$$\mathbb{X}_n = \frac{1}{\sqrt{N}} \begin{pmatrix} \mathbf{0}_{p,p} & \mathcal{X}_{N,p}^T \\ \mathcal{X}_{N,p} & \mathbf{0}_{N,N} \end{pmatrix} \text{ and } \mathbb{Y}_n = \frac{1}{\sqrt{N}} \begin{pmatrix} \mathbf{0}_{p,p} & \mathcal{G}_{N,p}^T \\ \mathcal{G}_{N,p} & \mathbf{0}_{N,N} \end{pmatrix} .$$

Indeed

$$S_{\mu_{\mathbf{B}_N}}(z) = z^{-1/2} \frac{n}{2p} S_{\mu_{\mathbb{X}_n}}(z^{1/2}) + \frac{N-p}{2pz}$$

Note that $\mathbb{X}_n := N^{-1/2}[x_{ij}^{(n)}]_{i,j=1}^n$ and $\mathbb{Y}_n := N^{-1/2}[y_{ij}^{(n)}]_{i,j=1}^n$ where

$$z_{ij}^{(n)} = \begin{cases} Z_{i-p,j} \mathbf{1}_{i \geq p+1} \mathbf{1}_{1 \leq j \leq p} & \text{if } 1 \leq j \leq i \leq n \\ z_{ji}^{(n)} & \text{if } 1 \leq i < j \leq n \end{cases} .$$

- The convergence (*) is handled via the Lindeberg method.

Preparatory materials to use Lindeberg's method

- Set $L_i^X = (x_{i1}^{(n)}, \dots, x_{ii}^{(n)})$ and $L_i^Y = (y_{i1}^{(n)}, \dots, y_{ii}^{(n)})$. Note that

$$S_{\mu_{\mathbb{X}_n}}(z) - S_{\mu_{\mathbb{Y}_n}}(z) = s(L_1^X, \dots, L_n^X) - s(L_1^Y, \dots, L_n^Y)$$

  where $s := s_z : \mathbb{R}^{n(n+1)/2} \to \mathbb{C}$. This function admits partial derivatives of all orders that are uniformly bounded (see Chatterjee (2006)).

Preparatory materials to use Lindeberg's method

- Set $L_i^X = (x_{i1}^{(n)}, \dots, x_{ii}^{(n)})$ and $L_i^Y = (y_{i1}^{(n)}, \dots, y_{ii}^{(n)})$. Note that

$$S_{\mu_{\mathbb{X}_n}}(z) - S_{\mu_{\mathbb{Y}_n}}(z) = s(L_1^X, \dots, L_n^X) - s(L_1^Y, \dots, L_n^Y)$$

  where $s := s_z : \mathbb{R}^{n(n+1)/2} \to \mathbb{C}$. This function admits partial derivatives of all orders that are uniformly bounded (see Chatterjee (2006)).

- Divide each row in big blocks of size m and small blocks of size q (the entries in these small blocks are replaced by 0). Hence

$$L_i \to \left(B_{i,1}, \mathbf{0}_q, B_{i,2}, \mathbf{0}_q, \dots, B_{i,k_i}, \mathbf{0}\right), \; k_i = [i/(m+q)]$$

  The asymptotics used will be $n \to \infty$ followed by $m \to \infty$ ($q = o(m)$).

Preparatory materials to use Lindeberg's method

- Set $L_i^X = (x_{i1}^{(n)}, \ldots, x_{ii}^{(n)})$ and $L_i^Y = (y_{i1}^{(n)}, \ldots, y_{ii}^{(n)})$. Note that

$$S_{\mu_{\mathbb{X}_n}}(z) - S_{\mu_{\mathbb{Y}_n}}(z) = s(L_1^X, \ldots, L_n^X) - s(L_1^Y, \ldots, L_n^Y)$$

  where $s := s_z : \mathbb{R}^{n(n+1)/2} \to \mathbb{C}$. This function admits partial derivatives of all orders that are uniformly bounded (see Chatterjee (2006)).

- Divide each row in big blocks of size m and small blocks of size q (the entries in these small blocks are replaced by 0). Hence

$$L_i \to \left(B_{i,1}, \mathbf{0}_q, B_{i,2}, \mathbf{0}_q, \ldots, B_{i,k_i}, \mathbf{0}\right), \ k_i = [i/(m+q)]$$

  The asymptotics used will be $n \to \infty$ followed by $m \to \infty$ ($q = o(m)$).

- Martingale approximation :

$$(B_{i,k}, \mathbf{0}) \to \widetilde{B}_{i,k} = (B_{i,k}, \mathbf{0}) - \mathbf{E}((B_{i,k}, \mathbf{0}) | \mathcal{F}_{i,k-1})$$

  where $\mathcal{F}_{i0} = \{\emptyset, \Omega\}$ and for $\ell \in \{1, \ldots, k_i\}$, $\mathcal{F}_{i\ell} = \sigma(B_{i,j}, j \leq \ell)$. Hence

$$L_i \to \widetilde{L}_i := \left(\widetilde{B}_{i,1}, \widetilde{B}_{i,2}, \ldots, \widetilde{B}_{i,k_i}\right), \ k_i = [i/(m+q)]$$

On the martingale approximations by blocks

How to control the expectations of the quantities

$\Delta_1 := s(L_1^X, \dots, L_n^X) - s(\widetilde{L}_1^X, \dots, \widetilde{L}_n^X)$ and $\Delta_2 := s(L_1^Y, \dots, L_n^Y) - s(\widetilde{L}_1^Y, \dots, \widetilde{L}_n^Y)$

On the martingale approximations by blocks

How to control the expectations of the quantities

$\Delta_1 := s(L_1^X, \dots, L_n^X) - s(\widetilde{L}_1^X, \dots, \widetilde{L}_n^X)$ and $\Delta_2 := s(L_1^Y, \dots, L_n^Y) - s(\widetilde{L}_1^Y, \dots, \widetilde{L}_n^Y)$

- Lemma (Götze *et al.* (2012)): Let **A** and **B** be two symmetric $n \times n$ matrices with real entries. Then, for any $z = x + \mathrm{i}y \in \mathbb{C} \backslash \mathbb{R}$,

$$|S_{\mu_{\mathbf{A}}}(z) - S_{\mu_{\mathbf{B}}}(z)| \leq \frac{1}{y^2\sqrt{n}} |\mathrm{Tr}(\mathbf{A} - \mathbf{B})^2|^{1/2} .$$

On the martingale approximations by blocks

How to control the expectations of the quantities

$\Delta_1 := s(L_1^X, \dots, L_n^X) - s(\widetilde{L}_1^X, \dots, \widetilde{L}_n^X)$ and $\Delta_2 := s(L_1^Y, \dots, L_n^Y) - s(\widetilde{L}_1^Y, \dots, \widetilde{L}_n^Y)$

- Lemma (Götze *et al.* (2012)): Let $\mathbf{A}$ and $\mathbf{B}$ be two symmetric $n \times n$ matrices with real entries. Then, for any $z = x + iy \in \mathbb{C} \backslash \mathbb{R}$,

$$|S_{\mu_{\mathbf{A}}}(z) - S_{\mu_{\mathbf{B}}}(z)| \leq \frac{1}{y^2 \sqrt{n}} |\mathrm{Tr}(\mathbf{A} - \mathbf{B})^2|^{1/2} .$$

- Therefore, recalling that $\mathcal{G}_k = \sigma(Y_\ell, \ell \leq k)$,

$$\big|\mathbf{E}\Delta_1\big|^2 \ll \big(\frac{q}{m} + \frac{q+m}{n}\big)\mathbf{E}(X_0^2) + \|\mathbf{E}(X_0|\mathcal{G}_{-q})\|_2^2$$

which converges to zero letting $n \to \infty$, followed by $m \to \infty$.

- Since (Y_ℓ) is a Gaussian process with the same cov. structure as (X_ℓ).

$$\big|\mathbf{E}\Delta_2\big|^2 \ll \big(\frac{q}{m} + \frac{q+m}{n}\big)\mathbf{E}(Y_0^2) + \|\mathbf{E}(Y_0|\sigma(Y_\ell, \ell \leq -q)\|_2^2$$

$$\ll \big(\frac{q}{m} + \frac{q+m}{n}\big)\mathbf{E}(X_0^2) + \|\mathbf{E}(X_0|\mathcal{G}_{-q})\|_2^2$$

Lindeberg's method by blocks (1)

Write

$$s(\widetilde{L}_1^X, \ldots, \widetilde{L}_n^X) - s(\widetilde{L}_1^Y, \ldots, \widetilde{L}_n^Y)$$
$$= \sum_{i=1}^{n} \sum_{k=1}^{k_i} \left(s_{i,k}(\widetilde{B}_{i,k}^X) - s_{i,k}(\widetilde{B}_{i,k}^Y) \right)$$

where $s_{i,k}(\widetilde{B}_{i,k})$ is the Stieltjes transform associated with the matrix

$$\begin{pmatrix}
\widetilde{B}^X_{1,k_1} & & & & & & & & & \\
\widetilde{B}^X_{1,k_2} & \widetilde{B}^X_{2,k_2} & & & & & & & & \\
. & . & . & & & & & & & \\
. & . & . & . & & & & & & \\
. & . & . & . & . & & & & & \\
\widetilde{B}^X_{i,1} & \widetilde{B}^X_{i,2} & \ldots & \widetilde{B}^X_{i,k-1} & \widetilde{B}_{i,k} & \widetilde{B}^Y_{i,k+1} & \ldots & \widetilde{B}^Y_{i,k_i} & & \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & . & \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & . & \\
\widetilde{B}^Y_{n,1} & \widetilde{B}^Y_{n,2} & \ldots & \widetilde{B}^Y_{n,k-1} & \widetilde{B}^Y_{n,k} & \widetilde{B}^Y_{n,k+1} & \ldots & \ldots & . & \widetilde{B}^Y_{n,k_n}
\end{pmatrix}$$

Lindeberg's method by blocks (2)

- To control $\sum_{i=1}^{n}\sum_{k=1}^{k_i}\mathbf{E}\big(s_{i,k}(\widetilde{B}^X_{i,k})-s_{i,k}(\widetilde{B}^Y_{i,k})\big)$, we write

$$\sum_{i=1}^{n}\sum_{k=1}^{k_i}\mathbf{E}\big(s_{i,k}(\widetilde{B}^X_{i,k})-s_{i,k}(\widetilde{B}^Y_{i,k})\big)$$

$$=\sum_{i=1}^{n}\sum_{k=1}^{k_i}\Big(\mathbf{E}\big(s_{i,k}(\widetilde{B}^X_{i,k})-s_{i,k}(\mathbf{0})\big)-\mathbf{E}\big(s_{i,k}(\widetilde{B}^Y_{i,k})-s_{i,k}(\mathbf{0})\big)\Big)$$

and we use Taylor's expansion at order three (so we need additionally to truncate the entries in the blocks $\widetilde{B}^X_{i,k}$). The terms of the first order are equal to zero due to the Martingale properties of the $\widetilde{B}^X_{i,k}$, the fact that the $\widetilde{B}^Y_{i,k}$'s are independent and the independence between (X_{ij}) and (Y_{ij}).

Lindeberg's method by blocks (2)

- To control $\sum_{i=1}^{n}\sum_{k=1}^{k_i}\mathbf{E}\big(s_{i,k}(\widetilde{B}^X_{i,k}) - s_{i,k}(\widetilde{B}^Y_{i,k})\big)$, we write

$$\sum_{i=1}^{n}\sum_{k=1}^{k_i}\mathbf{E}\big(s_{i,k}(\widetilde{B}^X_{i,k}) - s_{i,k}(\widetilde{B}^Y_{i,k})\big)$$
$$= \sum_{i=1}^{n}\sum_{k=1}^{k_i}\Big(\mathbf{E}\big(s_{i,k}(\widetilde{B}^X_{i,k}) - s_{i,k}(\mathbf{0})\big) - \mathbf{E}\big(s_{i,k}(\widetilde{B}^Y_{i,k}) - s_{i,k}(\mathbf{0})\big)\Big)$$

  and we use Taylor's expansion at order three (so we need additionally to truncate the entries in the blocks $\widetilde{B}^X_{i,k}$). The terms of the first order are equal to zero due to the Martingale properties of the $\widetilde{B}^X_{i,k}$, the fact that the $\widetilde{B}^Y_{i,k}$'s are independent and the independence between (X_{ij}) and (Y_{ij}).

- The main difficulty is to control suitably the terms of second order. Using only the fact that for any $z \in \mathbb{C}^+$

$$|\partial_{\mathbf{u}}\partial_{\mathbf{v}} s(\cdot)| \le \frac{c}{nN}$$

  does not lead to the desired result.

The Gaussian part

Let us prove that $\mathbf{E}S_{\mu_{\mathbf{G}_N}}(z) \to S(z)$.

- Let $(\xi_{ij})_{i,j\in\mathbb{Z}}$ be an iid Gaussian standard random field and let

$$\Gamma_p := \begin{pmatrix} c_0 & c_1 & \cdots & c_{p-1} \\ c_1 & c_0 & & c_{p-2} \\ \vdots & \vdots & \vdots & \vdots \\ c_{p-1} & c_{p-2} & \cdots & c_0 \end{pmatrix} \quad \text{where } c_k = \mathrm{Cov}(X_0, X_k)\,.$$

  We have $\mathbf{E}S_{\mu_{\mathbf{G}_N}}(z) = \mathbf{E}S_{\mu_{\mathbf{H}_N}}(z)$ where

$$\mathbf{H}_N = \frac{1}{N}\Gamma_p^{1/2}\Xi_{N,p}^T\Xi_{N,p}\Gamma_p^{1/2} \text{ with } \Xi_{N,p} = (\xi_{ij})_{1\le i\le N, 1\le j\le p}$$

- By Silverstein (95) + a version of Szegö's theorem for Toeplitz forms (see Trotter (84)), we get the result when the spectral density is in $\mathbb{L}^2$.

and when f is not in $\mathbb{L}^2$?

- Let
$$Y_{k\ell} = \sum_{j\in\mathbb{Z}} a_j \xi_{k,\ell-j} \text{ with } a_k = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{\mathrm{i}kx} \sqrt{f(x)} dx \,.$$
The process $(Y_{k\ell})_{k,\ell\in\mathbb{Z}}$ as the right covariance structure.

and when f is not in $\mathbb{L}^2$?

- Let
$$Y_{k\ell} = \sum_{j\in\mathbb{Z}} a_j \xi_{k,\ell-j} \text{ with } a_k = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{\mathrm{i}kx} \sqrt{f(x)} dx \,.$$
The process $(Y_{k\ell})_{k,\ell\in\mathbb{Z}}$ as the right covariance structure.
- Let b be a positive integer and let $f_b = f \wedge b$. Hence f_b is the spectral density on $[-\pi, \pi]$ of a $\mathbb{L}^2$-stationary process.

and when f is not in $\mathbb{L}^2$?

- Let
$$Y_{k\ell} = \sum_{j\in\mathbb{Z}} a_j \xi_{k,\ell-j} \text{ with } a_k = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{\mathrm{i}kx} \sqrt{f(x)} dx \,.$$
The process $(Y_{k\ell})_{k,\ell\in\mathbb{Z}}$ as the right covariance structure.
- Let b be a positive integer and let $f_b = f \wedge b$. Hence f_b is the spectral density on $[-\pi,\pi]$ of a $\mathbb{L}^2$-stationary process.
- Define
$$Z^b_{k\ell} = \sum_{j\in\mathbb{Z}} \tilde{a}_{j,b} \xi_{k,\ell-j} \text{ with } \tilde{a}_{k,b} = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{\mathrm{i}kx} \sqrt{f_b(x)} dx \,.$$

and when f is not in $\mathbb{L}^2$?

- Let
$$Y_{k\ell} = \sum_{j\in\mathbb{Z}} a_j \xi_{k,\ell-j} \text{ with } a_k = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{\mathrm{i}kx} \sqrt{f(x)} dx \,.$$
The process $(Y_{k\ell})_{k,\ell\in\mathbb{Z}}$ as the right covariance structure.
- Let b be a positive integer and let $f_b = f \wedge b$. Hence f_b is the spectral density on $[-\pi,\pi]$ of a $\mathbb{L}^2$-stationary process.
- Define
$$Z^b_{k\ell} = \sum_{j\in\mathbb{Z}} \tilde{a}_{j,b} \xi_{k,\ell-j} \text{ with } \tilde{a}_{k,b} = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{\mathrm{i}kx} \sqrt{f_b(x)} dx \,.$$
- $(Z^b_{\mathbf{u}})_{\mathbf{u}\in\mathbb{Z}^2}$ is a centered real-valued stationary Gaussian random field. In addition, for any fixed integer k, $(Z^b_{k\ell})_{\ell\in\mathbb{Z}}$ admits f_b as spectral density on $[-\pi,\pi]$. Let $\mathbf{G}^b_N$ be the Gram matrix associated with $(Z^b_{\mathbf{u}})_{\mathbf{u}\in\mathbb{Z}^2}$.

and when f is not in $\mathbb{L}^2$?

- Let
$$Y_{k\ell} = \sum_{j\in\mathbb{Z}} a_j \xi_{k,\ell-j} \text{ with } a_k = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{\mathrm{i}kx} \sqrt{f(x)} dx \,.$$
The process $(Y_{k\ell})_{k,\ell\in\mathbb{Z}}$ as the right covariance structure.
- Let b be a positive integer and let $f_b = f \wedge b$. Hence f_b is the spectral density on $[-\pi, \pi]$ of a $\mathbb{L}^2$-stationary process.
- Define
$$Z^b_{k\ell} = \sum_{j\in\mathbb{Z}} \tilde{a}_{j,b} \xi_{k,\ell-j} \text{ with } \tilde{a}_{k,b} = \frac{1}{\sqrt{2\pi}} \int_{-\pi}^{\pi} e^{\mathrm{i}kx} \sqrt{f_b(x)} dx \,.$$
- $(Z^b_{\mathbf{u}})_{\mathbf{u}\in\mathbb{Z}^2}$ is a centered real-valued stationary Gaussian random field. In addition, for any fixed integer k, $(Z^b_{k\ell})_{\ell\in\mathbb{Z}}$ admits f_b as spectral density on $[-\pi, \pi]$. Let $\mathbf{G}^b_N$ be the Gram matrix associated with $(Z^b_{\mathbf{u}})_{\mathbf{u}\in\mathbb{Z}^2}$.
- Since f_b is bounded, there exists a nonrandom p.m μ_b such that
$$\lim_{N\to\infty} d(F^{\mathbf{G}^b_N}, F^b) = 0 \text{ a.s.}$$
where d is the Lévy distance

- To control $d(F^{\mathbf{G}_N}, F^{\mathbf{G}_N^b})$, we use the following lemma

- To control $d(F^{\mathbf{G}_N}, F^{\mathbf{G}_N^b})$, we use the following lemma
- **Lemma :** Let $\mathbf{A}$ and $\mathbf{B}$ be two $p \times N$ matrices with real entries. Then

$$d^2(F^{\mathbf{A}\mathbf{A}^T}, F^{\mathbf{B}\mathbf{B}^T}) \leq \frac{\sqrt{2}}{p}\left[\mathrm{Tr}(\mathbf{A}\mathbf{A}^T + \mathbf{B}\mathbf{B}^T)\mathrm{Tr}((\mathbf{A}-\mathbf{B})(\mathbf{A}-\mathbf{B})^T)\right]^{1/2}.$$

- To control $d(F^{\mathbf{G}_N}, F^{\mathbf{G}_N^b})$, we use the following lemma
- **Lemma :** Let $\mathbf{A}$ and $\mathbf{B}$ be two $p \times N$ matrices with real entries. Then

$$d^2(F^{\mathbf{A}\mathbf{A}^T}, F^{\mathbf{B}\mathbf{B}^T}) \leq \frac{\sqrt{2}}{p}[\mathrm{Tr}(\mathbf{A}\mathbf{A}^T + \mathbf{B}\mathbf{B}^T)\mathrm{Tr}((\mathbf{A}-\mathbf{B})(\mathbf{A}-\mathbf{B})^T)]^{1/2}.$$

- The above lemma together with Cauchy-Schwarz's inequality and Parseval's identity give

$$\mathbf{E}d^2(F^{\mathbf{G}_N}, F^{\mathbf{G}_N^b}) \ll \Big(\int_{-\pi}^{\pi} f(x)dx\Big)^{1/2}\Big(\int_{-\pi}^{\pi} f(x)\,\mathbf{1}_{f(x)>b}dx\Big)^{1/2}$$

- To control $d(F^{\mathbf{G}_N}, F^{\mathbf{G}_N^b})$, we use the following lemma
- **Lemma :** Let **A** and **B** be two $p \times N$ matrices with real entries. Then

$$d^2(F^{\mathbf{A}\mathbf{A}^T}, F^{\mathbf{B}\mathbf{B}^T}) \leq \frac{\sqrt{2}}{p}[\mathrm{Tr}(\mathbf{A}\mathbf{A}^T + \mathbf{B}\mathbf{B}^T)\mathrm{Tr}((\mathbf{A}-\mathbf{B})(\mathbf{A}-\mathbf{B})^T)]^{1/2}.$$

- The above lemma together with Cauchy-Schwarz's inequality and Parseval's identity give

$$\mathbf{E}d^2(F^{\mathbf{G}_N}, F^{\mathbf{G}_N^b}) \ll \Big(\int_{-\pi}^{\pi} f(x)dx\Big)^{1/2}\Big(\int_{-\pi}^{\pi} f(x)\,\mathbf{1}_{f(x)>b}dx\Big)^{1/2}$$

- So finally $\lim_{b\to\infty}\limsup_{N\to\infty} d(F^{\mathbf{G}_N}, F^b) = 0$ in probability. In particular, F^b is Cauchy.

- To control $d(F^{\mathbf{G}_N}, F^{\mathbf{G}_N^b})$, we use the following lemma
- **Lemma :** Let **A** and **B** be two $p \times N$ matrices with real entries. Then

$$d^2(F^{\mathbf{A}\mathbf{A}^T}, F^{\mathbf{B}\mathbf{B}^T}) \leq \frac{\sqrt{2}}{p}[\mathrm{Tr}(\mathbf{A}\mathbf{A}^T + \mathbf{B}\mathbf{B}^T)\mathrm{Tr}((\mathbf{A}-\mathbf{B})(\mathbf{A}-\mathbf{B})^T)]^{1/2}.$$

- The above lemma together with Cauchy-Schwarz's inequality and Parseval's identity give

$$\mathbf{E}d^2(F^{\mathbf{G}_N}, F^{\mathbf{G}_N^b}) \ll \Big(\int_{-\pi}^{\pi} f(x)dx\Big)^{1/2}\Big(\int_{-\pi}^{\pi} f(x)\,\mathbf{1}_{f(x)>b}dx\Big)^{1/2}$$

- So finally $\lim_{b\to\infty}\limsup_{N\to\infty} d(F^{\mathbf{G}_N}, F^b) = 0$ in probability. In particular, F^b is Cauchy.
- Since the space of distribution functions endowed with d is complete there exists a nonrandom d.f. F such that $\lim_{b\to\infty} d(F^b, F) = 0$.

- To control $d(F^{\mathbf{G}_N}, F^{\mathbf{G}_N^b})$, we use the following lemma
- **Lemma :** Let $\mathbf{A}$ and $\mathbf{B}$ be two $p \times N$ matrices with real entries. Then

$$d^2(F^{\mathbf{A}\mathbf{A}^T}, F^{\mathbf{B}\mathbf{B}^T}) \leq \frac{\sqrt{2}}{p}[\mathrm{Tr}(\mathbf{A}\mathbf{A}^T + \mathbf{B}\mathbf{B}^T)\mathrm{Tr}((\mathbf{A}-\mathbf{B})(\mathbf{A}-\mathbf{B})^T)]^{1/2}.$$

- The above lemma together with Cauchy-Schwarz's inequality and Parseval's identity give

$$\mathbf{E}d^2(F^{\mathbf{G}_N}, F^{\mathbf{G}_N^b}) \ll \Big(\int_{-\pi}^{\pi} f(x)dx\Big)^{1/2}\Big(\int_{-\pi}^{\pi} f(x)\,\mathbf{1}_{f(x)>b}dx\Big)^{1/2}$$

- So finally $\lim_{b\to\infty}\limsup_{N\to\infty} d(F^{\mathbf{G}_N}, F^b) = 0$ in probability. In particular, F^b is Cauchy.
- Since the space of distribution functions endowed with d is complete there exists a nonrandom d.f. F such that $\lim_{b\to\infty} d(F^b, F) = 0$.
- By the continuity theorem, $S^b \to S$ and $S_{\mu_{\mathbf{G}_N}} \to S$ in probability.

- To control $d(F^{\mathbf{G}_N}, F^{\mathbf{G}_N^b})$, we use the following lemma
- **Lemma :** Let **A** and **B** be two $p \times N$ matrices with real entries. Then

$$d^2(F^{\mathbf{AA}^T}, F^{\mathbf{BB}^T}) \leq \frac{\sqrt{2}}{p}[\mathrm{Tr}(\mathbf{AA}^T + \mathbf{BB}^T)\mathrm{Tr}((\mathbf{A}-\mathbf{B})(\mathbf{A}-\mathbf{B})^T)]^{1/2}.$$

- The above lemma together with Cauchy-Schwarz's inequality and Parseval's identity give

$$\mathbf{E}d^2(F^{\mathbf{G}_N}, F^{\mathbf{G}_N^b}) \ll \Big(\int_{-\pi}^{\pi} f(x)dx\Big)^{1/2}\Big(\int_{-\pi}^{\pi} f(x)\,\mathbf{1}_{f(x)>b}dx\Big)^{1/2}$$

- So finally $\lim_{b\to\infty}\limsup_{N\to\infty} d(F^{\mathbf{G}_N}, F^b) = 0$ in probability. In particular, F^b is Cauchy.
- Since the space of distribution functions endowed with d is complete there exists a nonrandom d.f. F such that $\lim_{b\to\infty} d(F^b, F) = 0$.
- By the continuity theorem, $S^b \to S$ and $S_{\mu_{\mathbf{G}_N}} \to S$ in probability.
- To identify S, it suffices to take the limit of S^b.

Some concluding remarks

- Since the limiting spectral distribution of sample covariance matrices generated by stationary and regular processes in $\mathbb{L}^2$ always exists, looking only at the limit of the spectral distribution is not enough to distinguish short range dependent processes to those exhibiting long memory.

Some concluding remarks

- Since the limiting spectral distribution of sample covariance matrices generated by stationary and regular processes in $\mathbb{L}^2$ always exists, looking only at the limit of the spectral distribution is not enough to distinguish short range dependent processes to those exhibiting long memory.
- Looking at results of second order (like CLT for linear statistics of the eigenvalues) or at the asymptotic behavior of the largest eigenvalue could (may be!) allow this distinction (the normalizing sequences in the long memory setting would be probably different than in the short memory one).