LIMITING SPECTRUM OF AUTOCOVARIANCE MATRICES

FREE PROBABILITY IN TIME SERIES

ARUP BOSE
Joint work with Monika Bhattacharjee

Indian Statistical Institute
Kolkata

STOCHASTIC SYSTEMS AND APPLICATIONS
September 10, 2014

Outline

- Large dimensional time series model.
- Random matrices.
- All partitions. Classical independence.
- Non-crossing partitions. Free independence.
- Application to time series models.

Time series

$\varepsilon_{t,p}$'s: *i.i.d.* with mean 0 and variance-covariance matrix I_p.

$A_{j,p}^{(n)}, j \geq 0$: $p \times p$ *coefficient matrices* (parameters).

$n \to \infty$, $p = p(n) \to \infty$ such that $\frac{p}{n} \to y \in (0, \infty)$.

Observations: $X_{t,p}^{(n)}$, $1 \leq t \leq n$ satisfy

$$X_{t,p}^{(n)} = \sum_{j=0}^{q} A_{j,p}^{(n)} \varepsilon_{t-j,p} \; t, n \geq 1 \text{ (almost surely).} \tag{1}$$

Linear process (large dimensional version of linear processes such as ARMA etc.) q can also be ∞.

Autocovariance matrix sequence

- The sample autocovariance matrix of order i is defined as

$$\hat{\Gamma}_{i,p} := \frac{1}{n}\sum_{t=i+1}^{n} X_{t,p}X'_{(t-i),p}, \quad i = 1, 2, 3, \ldots (n-1).$$

- Symmetrized autocovariances:

$$\hat{\Gamma}_0, \quad \hat{\Gamma}_i\hat{\Gamma}'_i, \quad \hat{\Gamma}_i + \hat{\Gamma}'_i$$

Does there exist limiting spectral distribution (LSD) for the above symmetric matrices, more generally for any polynomial which is symmetrics?

How can the limits be described?

How to use them for statistical purposes?

Distribution, independence and moments

Assume all variables have compact support.

Moments

$$m_n = E(X^n)$$

determine the distribution (the probability measure) uniquely.

Moment generating function

$$M(z) = \sum m_n \frac{z^n}{n!}.$$

Cumulants

Cumulant generating function

$$\log M(z) = \sum k_n \frac{z^n}{n!}.$$

Example:

$$k_1 = m_1, \quad k_2 = m_2 - m_1^2.$$

Also

$$m_1 = k_1, \quad m_2 = k_2 + k_1^2.$$

Standard Gaussian law, $k_2(G) = 1$, other cumulants are zero. [Using this derive all the moments].

Constant function c, $k_1(c) = c$, other cumulants are zero.

In general what is the relation between moments and cumulants and can this relation be generalised?

All partitions

Fix n. Consider P, the set of ALL partitions of $\{1, 2, \ldots n\}$.

Define the REVERSE REFINEMENT PARTIAL ordering:

$1_n = \{1, 2, \ldots n\}$ is the largest partition

$0_n = \{\{1\}, \{2\}, ...\{n\}\}$ is the smallest.

Multiplicative extension

Consider any partition π. Consider any block V of π.

Extend the sequence k_n and m_n to k_π and m_π in a multiplicative way.

$$k_\pi = \prod_{V \in \pi} k_{|V|} \text{ and } m_\pi = \prod_{V \in \pi} m_{|V|}.$$

Then

$$m_n = m_{1_n} = \sum_{\pi \in P} k_\pi = \sum_{\pi \leq 1_n} k_\pi.$$

Indeed

$$m_\tau = \sum_{\pi \leq \tau} k_\pi \text{ for all } \tau.$$

What about the reverse relation?

Moments in terms of cumulants

P is a lattice (max and min operations are defined). Consider

$$P^{(2)} = \{(\pi, \sigma) : \pi \leq \sigma\}.$$

Define the IDENTITY FUNCTION

$$\delta : P^{(2)} \to R \text{ as } I_{\pi=\sigma}.$$

Define

$$\zeta(\pi, \sigma) = 1 \text{ for all } \pi \leq \sigma.$$

Then we can write the moment-cumulant formula as

$$m_\pi = \sum_{\sigma \leq \pi} k_\sigma \zeta(\sigma, \pi).$$

Mobius function

There exists $\mu(\pi, \sigma)$ on $P^{(2)}$ which is the INVERSE of ζ with respect to a convolution:

$$(\zeta * \mu)(\pi, \sigma) = \sum_{\pi \leq \tau \leq \sigma} \zeta(\pi, \tau)\mu(\tau, \sigma) = \delta(\pi, \sigma).$$

Equivalently,

$$\zeta * \mu = \delta,$$

or

$$\sum_{\pi \leq \tau \leq \sigma} \mu(\tau, \sigma) = I_{\pi = \sigma}.$$

μ is called the MOBIUS FUNCTION.

Cumulants in terms of moments

$$m = k * \zeta.$$

Applying Mobius inversion μ,

$$m * \zeta = k,$$

or

$$\sum_{\sigma \leq \pi} m_\sigma \mu(\sigma, \pi) = k_\pi.$$

In particular

$$k_n = k_{1_n} = \sum_\sigma m_\sigma \mu(\sigma, 1_n).$$

Mixed moments and mixed cumulants

$m_n(X_1, \ldots X_n) = E(X_1 \ldots X_n)$.

Extend multiplicatively.

Define cumulants by Mobius inversion function.

If $X_i = X$ for all i, we get back the earlier single sequence of moments and cumulants.

(Classical) independence, moments and cumulants

Random variables $X_1, \ldots, X_n$ are (classical) independent

if and only if ALL mixed moments factorize:

$$E(X_1^i X_2^j \ldots.) = E(X_1^i)E(X_2^j)\ldots.$$

if and only if ALL mixed cumulants are 0:

$k(X_i, X_j, \ldots) = 0$ whenever at least two indices are different.

Enter the non-commutative world

Random variables are elements of a C^* algebra where a linear function ϕ (called a STATE) is defined such that $\phi(1) = 1$. The first 1 is the unity of the algebra.

Examples:

All $n \times n$ matrices with constant entries and ϕ=normalized trace. The identity matrix is the unity 1.

All $n \times n$ matrices with random variable entries with ϕ=expected normalized trace.

Three types of independence

Suppose we wish to define a notion of independence between sub-algebras (NOT sigma-algebras). In a general sense there are only three possibilities:

Classical independence (a necessarily commutative notion).

Boolean independence (limited).

FREE INDEPENDENCE (non-commutative, rich).

Non-crossing partitions

Arrange the n points on a circle. Join the points that are in the same partition block.

If arcs of different partition blocks do not cross, then the partition is NON-CROSSING.

Let NC be the set of all non-crossing partitions.

Then NC inherits the ordering of P and it also has the lattice structure (closed under max and min operations).

Free cumulants

Now take variables $a_1, \ldots a_n$. Define

$$m_n(a_1 \ldots a_n) = \phi(a_1 \ldots a_n).$$

These may or may not define a probability distribution.

Extend in a multiplicative way (preserve the order of the non-commuting variables) on NC.

Mobius function exists on NC (always does on a POSET).

Define cumulants via this new Mobius function. Call them FREE CUMULANTS.

We can of course recover the moments from the cumulants.

Examples of free cumulants

1. Semi-circle law:

$$\frac{1}{2\pi}\sqrt{4-t^2}, \quad |t| \leq 2.$$

$k_2(s,s) = 1$. Other (free) cumulants are zero. Compare with standard Gaussian law: same thing happened with usual cumulants.

2. $\mu = \frac{1}{2}(\delta_{-1} + \delta_1)$. All odd cumulants are zero. Even (free) cumulants are

$$k_n(\mu, \ldots, \mu) = (-1)^{k-1} C_{k-1} \text{ if } n = 2k$$

where $C_j = \frac{1}{j+1}\binom{2j}{j}$ is the Catalan number.

FREE INDEPENDENCE

Declare variables $a_1, \ldots, a_n$ to be FREE INDEPENDENT *if and only if* all mixed free cumulants are zero.

Algebras (not sigma-algebras) of free variables are free.

If a, b are free, then cumulants add:

$$k_n(a+b, \ldots a+b) = k_n(a, \ldots, a) + k_n(b, \ldots, b).$$

The free ADDITIVE CONVOLUTION $\mu \boxplus \nu$, of two probability distributions μ and ν is again a probability distribution. Its free cumulants are the sum of the free cumulants of μ and ν. If a and b are free with distributions μ and ν then the distribution of $a+b$ is $\mu \boxplus \nu$.

Free product convolution is a bit more tricky..

Example of free additive convolution

1. $\mu = \nu = \frac{1}{2}(\delta_{-1} + \delta_1)$. Then $\gamma = \mu \boxplus \nu$ is the arc sine law with density

$$\frac{1}{\pi\sqrt{4 - t^2}}, \quad |t| \leq 2.$$

So if a and b are two symmetric Bernoulii and are free then the distribution of $a + b$ is arc sine. In general, free additive convolution of discrete measures can be continuous.

2. s_1, s_2 two free semi-circle variables (such a combination does exist). Then $\frac{s_1+s_2}{\sqrt{2}}$ is again semi circle. So semi-circle (which has compact support) is free infinitely divisible.

Asymptotic free

$(\mathcal{A}_N, \phi_N)$. Say that $(a_{1,N}, \ldots, a_{k,N})$ converges JOINTLY *if and only if* all their moments converge *if and only if* all their cumulants converge.

Define a limit algebra of k indeterminates and a state via this limit ϕ:

$$\phi(a_1, \ldots, a_k) = \lim_{N \to \infty} \phi_N(a_{1,N}, \ldots, a_{k,N}).$$

$(a_{1,N}, \ldots, a_{k,N})$ are said to be ASYMPTOTICALLY free if the corresponding limits are free (with respect to ϕ).

Wigner matrices and asymptotic freeness

Wigner matrix: $N \times N$ symmetric matrix with (clasical) independent and identically distributed variance one entries. As $N \to \infty$

1 (a) Its (scaled) spectral distribution (random distribution of eigen values) converges weakly almost surely to the semicircle law. Such limits are usually called LSD.

(b) If all moments are assumed finite then all its moments converge (almost surely) to the moments of the semi-circle law.

2. Independent Wigner matrices converge jointly and are asymptotically free (with marginals as semi-circle).

3. Such matrices are also free of non-random norm bounded matrices (which may not be free between themselves) which converge jointly.

Recall: the sample autocovariance matrices:

$$\hat{\Gamma}_{i,p} := \frac{1}{n}\sum_{t=i+1}^{n} X_{t,p}X'_{(t-i),p} = \sum_{j=0}^{q}\sum_{j'=0}^{q} A_j\hat{\Gamma}_{j'-j+i}(\varepsilon)A'_{j'}.$$

Note that approximately

$$\hat{\Gamma}_k(\varepsilon) = WP_kW$$

where W is a Wigner matrix and P_k is a non-random matrix whose kth diagonal equals 1.

Let $(a_i, 1 \leq i \leq q)$ = joint limit of $(A_i, 1 \leq i \leq q)$,

Let $(c_i, 1 \leq i \leq q)$ = joint limit of P_i.

Let s be standard semi-circular and

$(a_i, a_i*, 1 \leq i \leq q)$ s and $(c_i, c_i*, 1 \leq i \leq q)$ be free.

Theorem

Define (involves non-cummutative products)

$$\gamma_{i,q} := \sum_{j=0}^{q} \sum_{j'=0}^{q} a_j \; s \; c_{j'-j+i} \; s \; a_{j'}^{*} \quad \forall i = 0, 1, 2, \ldots$$

Now we are ready to write our main result.

Theorem

• Under finiteness of all moments, there is joint convergence of the random and non-random matrices. In particular $\hat{\Gamma}_i$ converge to γ_i (other than some scaling and masses at zero).

• Under enough moments, LSD for any symmetric polynomial, say $P(\hat{\Gamma}_i, \hat{\Gamma}_i^* : \; i \geq 0)$ exists.

No closed form expression for the distribution, except in special cases. Can calculate moments of any order recursively.

Suppose $X_t = \varepsilon_t \; \forall t$. Then

• $\hat{\Gamma}_0(\varepsilon)$ is the usual sample variance-covariance matrix and its LSD is a Marčenko-Pastur law.

• The LSD of $\hat{\Gamma}_i(\varepsilon)\hat{\Gamma}_i(\varepsilon)'$ is the Bessel$(2, y^{-1})$ law.

Jin, Wang, Bai, Nair and Harding (2014): existence of LSD of $\frac{1}{2}(\hat{\Gamma}_i + \hat{\Gamma}_i')$ when $X_t = \varepsilon_t \ \forall t$, via Stieltjes transform.

Liu, Ahn and Paul (2013): LSD of $\hat{\Gamma}_0$ and $\frac{1}{2}(\hat{\Gamma}_i + \hat{\Gamma}_i')$ for MA(q) under strong conditions on the coefficient matrices via Stieltjes transform.

Jin, Wang, Miao and Lo Huang (2009): existence of LSD for sample covariance matrix of X_t when it has i.i.d. rows and each row is a VARMA.

All these results become special cases of our result.

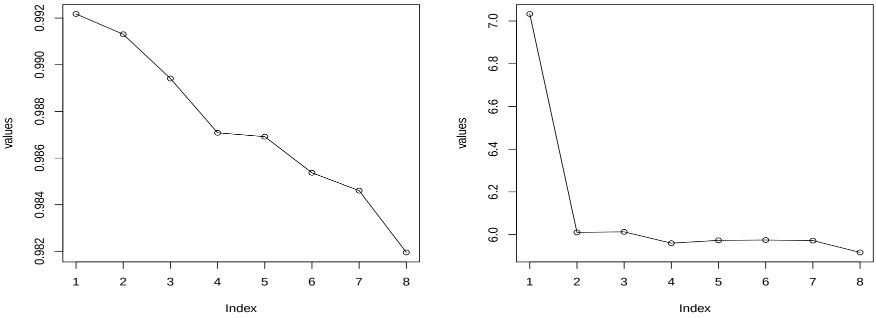
Diagnosis

When X_t is MA(q), the LSD of $\hat{\Gamma}_i\hat{\Gamma}_i'$ are identical for $i > q$ and are different for $0 \leq i \leq q$.
Consider

$$X_t = \sum_{k=0}^{q} \varepsilon_{t-k}, \quad q = 0, 1.$$

We also let $y = 1$. For each fixed $q = 0, 1$ and $r = 1, 2, 3$, we plot the r-th order moments of $\hat{\Gamma}_i\hat{\Gamma}_i'$ for $i = 1, 2, \ldots, 8$. Observe that, for each $r = 1, 2, 3$, the r-th order moments are more or less same for each $i = q + 1, q + 2, \ldots, 8$ under MA(q) process, $q = 0, 1$.

Also, the 1st, 2nd and 3rd order moments when X_t is MA(0), are near about 1, 3 and 12 respectively, which are indeed the 1st , 2nd and 3rd order moments of Bessel(2,1) law.



*F*igure 1: MA(0) and MA(1) processes: 1st moment of $\hat{\Gamma}_i\hat{\Gamma}_i'$

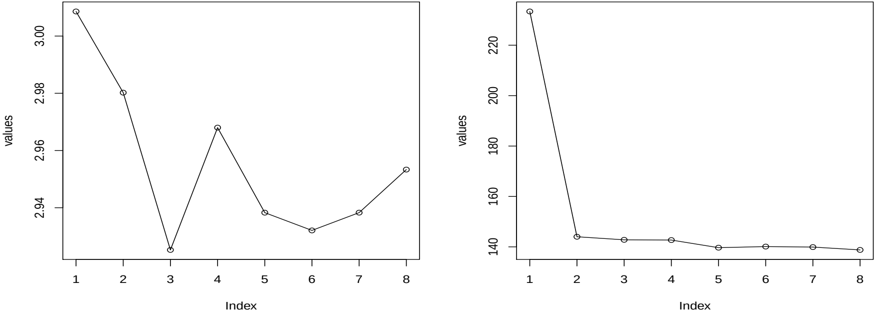
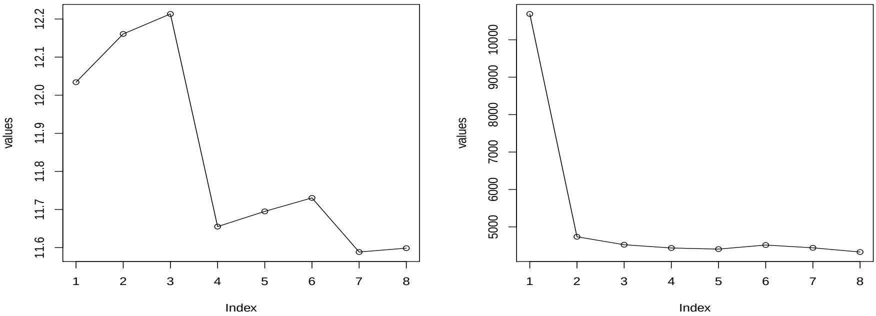


Figure 2: MA(0) and MA(1) processes: 2nd moment of $\hat{\Gamma}_i \hat{\Gamma}_i'$



*F*igure 3: MA(0) and MA(1) processes: 3rd moment of $\hat{\Gamma}_i \hat{\Gamma}_i'$

Bessel law

Moments are given by:

$$\beta_h = \sum_{k=1}^{h} \frac{1}{k} \binom{h-1}{k-1} \binom{2h}{k-1} y^{-k}.$$

Jin, Baisuo; Wang, Chen; Bai, Z. D.; Nair, K. Krishnan and Harding, Matthew (2014). Limiting spectral distribution of a symmetrized auto-cross covariance matrix, *Ann. Appl. Probab.*, 1199–1225.

Jin, Baisuo; Wang, Cheng; Miao, Baiqi and Lo Huang, Mong-Na (2009). Limiting spectral distribution of large-dimensional sample covariance matrices generated by VARMA, *J. Multivariate Anal.*, 2112–2125.

Liu, H.; Aue, A. and Paul, D. (2013). On the Marčenko-Pastur law for linear time series, *Annals of Statist.*, 199–227.

Nica, A. and Speicher, R. (2006). *Lectures on the combinatorics of free probability*, Cambridge University Press, Cambridge, UK.