Lecture 24: May 25, 2011

Convergence in Non-Commutative Distribution

Set $A^N = \{a_i^N\}_{i\in I}$ be a collection of random variables for each N.

Recall that the distribution μ_{A^N} of this family is just the collection of all joint moments in them. This collection can be thought of as a linear functional on $\mathbb{C}[X_i]_{i\in I}$, the algebra of non-commutative polynomials in the indeterminates X_i.

The key feature here is that μ_{A^N} has no access to the NCPS in which the elements of A^N live. Thus, we may as well take them in different spaces.

Def: Let $(\mathcal{A}_N, \varphi_N)$ be NCPS's for $N\in\mathbb{N}$. For a fixed index set I, let $A^N = \{a_i^N\}_{i\in I}$ be a set of random variables in $\mathcal{A}_N$. Let μ_N be the distribution of A^N: $\mu_N \in \mathbb{C}[X_i : i\in I]'$, defined by

$$\mu_N(P) = \varphi_N(P(a_i^N : i\in I))$$

(E.g. $\mu_N(X_1X_2X_1X_3) = \varphi_N(a_1^N a_2^N a_1^N a_3^N)$)

Let $(\mathcal{A},\varphi)$ be a NCPS, with $A = \{a_i\}_{i\in I} \subseteq \mathcal{A}$. Then $\mu = \mu_A \in \mathbb{C}[X_i : i\in I]'$ as well.

Say $\mu_N \to \mu$ (ie. $\{a_i^N\}_{i\in I}$ converges in distribution to $\{a_i\}_{i\in I}$, $A^N \xrightarrow{\mathcal{D}} A$) if

$$\forall P\in \mathbb{C}[X_i : i\in I] \quad \mu_N(P) \to \mu(P) \text{ as } N\to\infty.$$

I.e. this is convergence of all joint moments separately — what we might call "formal convergence".

Eg. if $A^N = \{x_N\}$, $A = \{x\}$, then $x_N \xrightarrow{\mathcal{D}} x$ means $\varphi_N(x_N^k) \to \varphi(x_N^k)$ $\forall k\geq 1$.

This is weaker than classical convergence in distribution, which insists on some uniformity in the convergence among all moments. (For example, if $\exists R>0$ s.t. $|\varphi(x_N^k)| \leq R^k \ \forall N$, then we actually $\mu_N \rightarrow \mu$ in the classical sense from ere.)

Let's take a look now at an important example. Let

$$\mathcal{M}_N = \bigcap_{p\geq 1} L^p(\Omega, \mathcal{F}, \mathbb{P}; M_N(\mathbb{C}))$$

the set of random $N\times N$ matrices all of whose entries are in $L^p(\Omega,\mathcal{F},\mathbb{P})$ for all $p\geq 1$ – i.e.

$$\mathbb{E}(|X_{ij}|^p) < \infty \quad \forall p\geq 1,\ 1\leq i,j\leq N$$

finite moments of all orders. This is an algebra, by Hölder's inequality and Minkowski's inequality.

Let $\varphi_N(X) = \frac{1}{N}\mathbb{E}\,\mathrm{Tr}\,X$. This is evidently a state ($\mathbb{E}\,\mathrm{Tr}$ is linear and well-defined on $\mathcal{M}_N$ since $X\in\mathcal{M}_N$ has $\mathbb{E}(|X_{ij}|)<\infty\ \forall i,j$). Moreover

$$\varphi_N(XX^*) = \frac{1}{N}\mathbb{E}(\underbrace{\mathrm{Tr}\,XX^*}_{=\sum_{i,j=1}^N |X_{ij}|^2}) = \frac{1}{N}\sum_{i,j=1}^N \mathbb{E}(|X_{ij}|^2) \geq 0$$

and only equals 0 if $\mathbb{E}(|X_{ij}|^2)=0\ \forall i,j$ meaning that $X_{ij}=0$ in L^2 so $X=0$ in $\mathcal{M}_N$. Thus

$$(\mathcal{M}_N, \varphi_N) \text{ is a NCPS.}$$

Now we consider a very special r.v. in this NCPS. On $(\Omega,\mathcal{F},\mathbb{P})$, select random variables

$$(G_j^0)_{1\leq j\leq N} \quad (G_{jk}^1)_{1\leq j<k\leq N} \quad (G_{jk}^2)_{1\leq j<k\leq N}$$

with all N^2 variables iid $N(0,1)$.

Construct from them a matrix $X = X_N \in M_N$ as follows:

$$1 \leq j \leq N: \quad X_{jj} = \alpha G^0_j$$

$$1 \leq j < k \leq N: \quad \begin{aligned} X_{jk} &= \frac{\alpha}{\sqrt{2}}\left(G^1_{jk} + i G^2_{jk}\right) \\ X_{kj} &= \frac{\alpha}{\sqrt{2}}\left(G^1_{jk} - i G^2_{jk}\right) \end{aligned}$$

The parameter $\alpha > 0$ will be chosen later: note, we can let it depend on N.

This random matrix is called a GUE_N: Gaussian Unitary Ensemble. (We will explain what "unitary" is doing in there later.) We can easily compute the covariance of any two entries:

$\mathbb{E}(X_{ij} X_{kl})$ first take the case of two upper-Δ entries:

$$= \mathbb{E}\left[\frac{\alpha}{\sqrt{2}}\left(G^1_{ij} + i G^2_{ij}\right)\frac{\alpha}{\sqrt{2}}\left(G^1_{kl} + i G^2_{kl}\right)\right] \quad (i<j,\ k<l)$$

$$= \frac{\alpha^2}{2}\left[\underbrace{\mathbb{E}(G^1_{ij} G^1_{kl})}_{=0 \text{ unless } (i,j)=(k,l)} + i\,\underbrace{\mathbb{E}(G^1_{ij} G^2_{kl})}_{0} + i\,\underbrace{\mathbb{E}(G^2_{ij} G^1_{kl})}_{0} - \underbrace{\mathbb{E}(G^2_{ij} G^2_{kl})}_{=0 \text{ unless } (i,j)=(k,l)}\right]$$

always, b/c G^1, G^2 indep.

So get $\frac{\alpha^2}{2}\delta_{ik}\delta_{jl}\left[\mathbb{E}\left((G^1_{ij})^2\right) - \mathbb{E}\left((G^2_{ij})^2\right)\right] = 0$.

Now take one upper-Δ, the other lower-Δ:

$$\mathbb{E}\left[\frac{\alpha}{\sqrt{2}}\left(G^1_{ij} + i G^2_{ij}\right)\frac{\alpha}{\sqrt{2}}\left(G^1_{lk} - i G^2_{lk}\right)\right] \quad i<j,\ k<l$$

$$= \frac{\alpha^2}{2}\left[\underbrace{\mathbb{E}(G^1_{ij} G^1_{lk})}_{\delta_{il}\delta_{jk}} - i\,\underbrace{\mathbb{E}(G^1_{ij} G^2_{lk})}_{0} + i\,\underbrace{\mathbb{E}(G^2_{ij} G^1_{lk})}_{0} + \underbrace{\mathbb{E}(G^2_{ij} G^2_{lk})}_{\delta_{il}\delta_{jk}}\right]$$

$$= \alpha^2 \delta_{il}\delta_{jk}.$$

Indeed, checking all other cases, we have

$$\mathbb{E}(X_{ij}X_{kl}) = \alpha^2 \delta_{il}\delta_{jk} \qquad 1 \le i,j,k,l \le N.$$

In other words, $\mathbb{E}(|X_{ij}|^2) = \mathbb{E}(X_{ij}\overline{X_{ij}}) = \mathbb{E}(X_{ij}X_{ji}) = \alpha^2$ for all i,j; all other covariances are 0.

Thus we can compute the first two moments of $X_N \in (M_N, \varphi_N)$:

$$\varphi(X_N) = \frac{1}{N}\mathbb{E}\,\mathrm{Tr}(X_N) = \frac{1}{N}\sum_{j=1}^{N}\mathbb{E}(G^{e}_{ii}) = 0.$$

$$\varphi(X_N^2) = \frac{1}{N}\mathbb{E}\,\mathrm{Tr}(X_N^2) = \frac{1}{N}\sum_{j=1}^{N}\mathbb{E}([X_N^2]_{jj})$$

$$= \frac{1}{N}\sum_{j,k=1}^{N}\mathbb{E}((X_N)_{jk}(X_N)_{kj})$$

$$= \frac{1}{N}\sum_{j,k=1}^{N}\mathbb{E}(|X_{jk}|^2) = \frac{1}{N}\cdot N^2\alpha^2$$

$$= N\alpha^2.$$

For standard normalization, this suggests we set $\alpha = 1/\sqrt{N}$, which we now promptly do.

We would like to compute higher moments of X_N. When we do, expectations of products of more than two guassians will come into play.

The collection $\{\mathrm{Re}\,X_{ij}, \mathrm{Im}\,X_{ij}\}_{1 \le i \le N}$ form a Gaussian random vector.

Def: A random vector $X \in \mathbb{R}^n$ is Gaussian if $\exists$ pos. def. matrix $C \in M_n(\mathbb{R})$ s.t.

$$\mu_{\underline{X}}(dt) = (2\pi)^{-n/2}(\det C)^{-1/2}\, e^{-\frac{1}{2}\langle t, C^{-1}t\rangle}\, d^n t.$$

The matrix C is the covariance matrix – i.e. $C_{ij} = \mathbb{E}(X_i X_j)$. The real/imaginary parts of the entries of our X_N have such a law, as we will discuss soon.

Wick's Formula: Let $\underline{X}$ be a Gaussian random vector in $\mathbb{R}^n$. Let $i_1, \ldots, i_k \in [n]$. Then

$$\mathbb{E}(X_{i_1} \cdots X_{i_k}) = \sum_{\pi \in \mathcal{P}_2(k)} \prod_{(r,s) \in \pi} \mathbb{E}(X_{i_r} X_{i_s}).$$

Pf. Both sides of the formula are multi-linear in the k variables. So it suffices to prove the formula holds for $\underline{Y}$ in place of $\underline{X}$ where $\underline{Y} = T(\underline{X})$ for some invertible linear transformation T of $\mathbb{R}^n$.

In particular, diagonalize $C = U^T \Lambda U$. Let $\underline{Y} = \Lambda U \underline{X}$. Simple calculation shows that

$$\mu_{\underline{Y}}(dt) = (2\pi)^{-n/2} e^{-|t|^2/2} dt.$$

In other words, $Y_1, \ldots, Y_n$ are iid $N(0,1)$ random variables. So it suffices to prove Wick's formula in this case.

Let $i_1, \ldots, i_k \in [n]$. Let $\tau \in \mathcal{P}(k)$ be the partition this multi-index determines (so $l \sim_\tau m$ iff $i_l = i_m$). Then if $\tau = \{B_1, \ldots, B_r\}$, and the index j_m represents B_m)

$$\mathbb{E}(Y_{i_1} \cdots Y_{i_k}) = \mathbb{E}(Y_{j_1}^{|B_1|} \cdots Y_{j_r}^{|B_r|}) = \mathbb{E}(Y_{j_1}^{|B_1|}) \cdots \mathbb{E}(Y_{j_r}^{|B_r|})$$

by independence.

As calculated earlier, $\mathbb{E}(Y_j^n) = \begin{cases} 0, & n \text{ odd} \\ (n-1)!! & n \text{ even} \end{cases} = \#\mathcal{P}_2(n)$.

So $\mathbb{E}(Y_{i_1} \cdots Y_{i_k}) = \#\mathcal{P}_2(|B_1|) \cdots \#\mathcal{P}_2(|B_r|)$. On the other hand

$$\prod_{(a,b) \in \pi} \mathbb{E}(Y_{i_a} Y_{i_b}) \neq 0 \text{ iff } i_a = i_b \ \forall (a,b) \in \pi$$

i.e. $a \sim_\pi b \Rightarrow a \sim_\tau b$

i.e. $\pi \leq \tau$.

In this case the product is $\prod \mathbb{E}(Y_j^2) = 1$. Thus

$$\sum_{\pi \in \mathcal{P}_2(k)} \prod_{(a,b) \in \pi} \mathbb{E}(Y_{i_a} Y_{i_b}) = \#\{\pi \in \mathcal{P}_2(k) : \pi \leq \tau\}.$$

It is easy to describe the set $\{\pi \in \mathcal{P}_2(k) : \pi \leq \tau\}$, as there is no non-crossing condition to worry about. Each block in π must be contained in some B_m. Thus

$$\{\pi \in \mathcal{P}_2(k) : \pi \leq \tau\} = \mathcal{P}_2(B_1) \times \cdots \times \mathcal{P}_2(B_r).$$

Thus $\displaystyle\sum_{\pi \in \mathcal{P}_2(k)} \prod_{(a,b) \in \pi} \mathbb{E}(Y_{i_a} Y_{i_b}) = \# \mathcal{P}_2(B_1) \times \cdots \times \mathcal{P}_2(B_r)$

$$= \# \mathcal{P}_2(|B_1|) \cdots \# \mathcal{P}_2(|B_r|).$$

///

Remark: We just proved Wick's formula for a (jointly) Gaussian real vector. But, again by the multi-linearity of the formula, if $Z_1, \ldots, Z_n$ are such that

$\mathrm{Re}\, Z_1, \ldots, \mathrm{Re}\, Z_n, \mathrm{Im}\, Z_1, \ldots, \mathrm{Im}\, Z_n$ form a real Gaussian vector,

then Wick's formula will hold with Z_i's in place of X_i's. By the same reasoning, we could use any combination of Z_i and $\bar{Z}_i$, since any such n-vector is a linear transformation of the $2n$-vector $\mathrm{Re}\, Z_1, \ldots, \mathrm{Im}\, Z_n$.

Exercise: Let X_N be a GUE_N (with parameter α). Calculate the covariance matrix $(2N^2 \times 2N^2)$, i.e.

$$C = \begin{bmatrix} C^1 & C^3 \\ C^3 & C^2 \end{bmatrix} \qquad \begin{aligned} C^1_{ij,kl} &= \mathbb{E}(\mathrm{Re}\, X_{ij}\, \mathrm{Re}\, X_{kl}) \\ C^2_{ij,kl} &= \mathbb{E}(\mathrm{Im}\, X_{ij}\, \mathrm{Im}\, X_{kl}) \\ C^3_{ij,kl} &= \mathbb{E}(\mathrm{Re}\, X_{ij}\, \mathrm{Im}\, X_{kl}) \end{aligned}$$

Verify that C is positive definite, so that the entries of X_N form a jointly Gaussian family.

Remark: Wick's formula is really the classical moment-cumulant formula for Gaussians

$$\mathbb{E}(X_{i_1} \cdots X_{i_k}) = \sum_{\pi \in \mathcal{P}(k)} c_\pi(X_{i_1}, \ldots, X_{i_k}).$$

Wick tells us that $c_\pi = 0$ if $\pi \notin \mathcal{P}_2$.

We already knew this for a single Gaussian; we've shown it's true for any Gaussian family.

Semicircular Systems & Families

Let $(\mathcal{A}, \varphi)$ be a NCPS. A collection $\{s_1, \ldots, s_n\}$ in $\mathcal{A}$ is called a <u>semicircular system</u> if each s_i is a standard semicircular r.v. and $s_1, \ldots, s_n$ are free. More generally, it is called a semicircular <u>family</u> if each s_i is semicircular and the joint <u>free</u> cumulants $\kappa_\pi(s_{i_1}, \ldots, s_{i_k}) = 0$ if $\pi \notin NC_2(k)$.
I.e.

$$\varphi(s_{i_1} \cdots s_{i_k}) = \sum_{\pi \in NC_2(k)} \kappa_\pi(s_{i_1} \cdots s_{i_k})$$

$$= \sum_{\pi \in NC_2(k)} \prod_{(a,b) \in \pi} \varphi(s_{i_a} s_{i_b}).$$

That is: "semicircular family" is <u>defined</u> by a free Wick's formula.

<u>Exercise</u>: Let $\{s_1, \ldots, s_n\}$ be a semicircular family. Let $C \in M_n(\mathbb{R})$ be its covariance matrix: $C_{ij} = \varphi(s_i s_j)$. Show that $s_1, \ldots, s_n$ are free iff C is diagonal; they are a semicircular system iff $C = I$.