Random matrix theory: lecture 22

Free probability (cont'd) [ref: Voiculescu]

Reminder: non-commutative probability

- $\mathcal{A}$ algebra with unit element I and involution $A \mapsto A^T$.
- $\varphi : \mathcal{A} \to \mathbb{R}$ linear and positive application s.t. $\varphi(I) = 1$

Example:

$\mathcal{A} = \{ n \times n$ real (deterministic) matrices $\}$

Let $\rho \in \mathcal{A}$ be s.t. $\rho \geq 0$ and $Tr(\rho) = 1$ ("density matrix")

$\varphi(A) := Tr(\rho A)$

sub-examples: $\rho = \begin{pmatrix} 1 & & 0 \\ & 0 & \\ 0 & & \ddots & \\ & & & 0 \end{pmatrix} \Rightarrow \varphi(A) = a_{11}$

$\rho = \frac{1}{n} I \Rightarrow \varphi(A) = \frac{1}{n} Tr A$

Terminology:

- elements of $\mathcal{A}$ are non-commutative random variables
- $\varphi(A)$ is the expectation of A
- the distribution of A is specified

  through its moments $m_k^A = \varphi(A^k)$, $k \geq 0$

  (NB: in the present framework, these always exist)

Note that in general, $\varphi(ABAB) \neq \varphi(A^2B^2)$.

Def: classical independence

Two non-commutative random variables A and B are classically independent if:

- they commute, i.e. $AB = BA$
- $\varphi(A^k B^\ell) = \varphi(A^k)\varphi(B^\ell) \quad \forall k, \ell \geq 0$

Def: free independence

Two non-commutative random variables A and B are freely independent if for any $k \geq 0$ and any polynomials $P_1 .. P_m$, $Q_1 .. Q_m$ such that $\varphi(P_j(A)) = \varphi(Q_j(A)) = 0 \quad \forall 1 \leq j \leq m$, we have $\varphi(P_1(A)Q_1(B)P_2(A)Q_2(B) \dots P_m(A)Q_m(B)) = 0$

[NB: $P_1(A)$ or $Q_m(B)$ might be replaced by I also]

Example: if A, B are freely independent, then

$$\varphi((A-\varphi(A)I)(B-\varphi(B)I)(A-\varphi(A)I)(B-\varphi(B)I)) = 0$$

since $\varphi(A-\varphi(A)I) = \varphi(A) - \varphi(A)\underbrace{\varphi(I)}_{=1} = 0$, and

similarly: $\varphi(B-\varphi(B)I) = 0$.

We will see below that the notion of free independence is too restrictive in the context of classical probability (i.e. commuting random variables).

Lemma 1

If A and B commute and are freely independent, then they are classically independent.

Proof:

Consider $m=1$ & $P_1(A)=A^k-\varphi(A^k)I$, $Q_1(B)=B^\ell-\varphi(B^\ell)I$

A,B free, $\varphi(P_1(A))=\varphi(Q_1(B))=0 \Rightarrow \varphi(P_1(A)Q_1(B))=0$

i.e. $\varphi((A^k-\varphi(A^k)I)(B^\ell-\varphi(B^\ell)I))=0$

i.e. $\varphi(A^kB^\ell)=\varphi(A^k)\varphi(B^\ell) \quad \forall k,\ell\geq 0$ #

Lemma 2

If A and B are freely independent, then

$$\varphi(A^2B^2)-\varphi(ABAB)=(\varphi(A^2)-\varphi(A)^2)(\varphi(B^2)-\varphi(B)^2)$$

Proof:

- A,B free $\Rightarrow \varphi((A^2-\varphi(A^2)I)(B^2-\varphi(B^2)I))=0$

  i.e. $\varphi(A^2B^2)=\varphi(A^2)\varphi(B^2)$

- define now $A_0=A-\varphi(A)I$, $B_0=B-\varphi(B)I$: $\varphi(A_0)=\varphi(B_0)=0$

  so $\underbrace{\varphi(A_0B_0A_0B_0)}=0$

$$=\varphi(AB_0A_0B)-\varphi(A)\underbrace{\varphi(B_0A_0B_0)}_{=0}-\underbrace{\varphi(A_0B_0A_0)}_{=0}\varphi(B)$$

$$-\ \varphi(A)\underbrace{\varphi(B_0A_0)}_{=0}\varphi(B)$$

In turn,

$$0 = \varphi(A B_0 A_0 B) = \varphi(ABAB) - \overbrace{\varphi(A^2B)}^{=\varphi(A^2)\varphi(B)}\varphi(B) - \overbrace{\varphi(AB^2)}^{=\varphi(A)\varphi(B^2)}\varphi(A)$$
$$+ \varphi(B)\underbrace{\varphi(AB)}_{=\varphi(B)\varphi(A)}\varphi(A)$$

So $\varphi(ABAB) = \varphi(A^2)\varphi(B)^2 + \varphi(A)^2\varphi(B^2) - \varphi(A)^2\varphi(B)^2$

and $\varphi(ABAB) - \varphi(A^2B^2) = (\varphi(A^2) - \varphi(A)^2)(\varphi(B^2) - \varphi(B)^2)$ #

Corollary 1

If A and B commute and are freely independent, then $0 = \varphi(A^2B^2) - \varphi(ABAB) = (\varphi(A^2) - \varphi(A)^2)(\varphi(B^2) - \varphi(B)^2)$ so either $\varphi(A^2) - \varphi(A)^2 = 0$ or $\varphi(B^2) - \varphi(B)^2 = 0$.

Corollary 2

× If A and B are classical randam variables and are freely independant, then either $Var(A) = \varphi(A^2) - \varphi(A)^2 = 0$ or $Var(B) = \varphi(B^2) - \varphi(B)^2 = 0$ i.e. one of the two randam variables is a constant. In classical probability, there are therefore only trivial examples of free randam variables.

Lemma 3

If A and B are freely independant, then the sequence $(\varphi((A+B)^k), k \geq 0)$ is entirely determined by the two sequences $(\varphi(A^k), k \geq 0)$ and $(\varphi(B^k), k \geq 0)$, i.e. the distribution of $A+B$ is determined by the distribution of A and the distribution of B separately.

Remark: This is to relate to what we know in classical probability: if two random variables X and Y are independent, then the distribution of $X+Y$ is determined by the distributions of X and Y; more precisely, the distribution of $X+Y$ is the (classical) convolution:

$$\mu_{X+Y} = \mu_X * \mu_Y$$

Proof of the above lemma: (main idea)

The lemma can be shown by induction on k. The main idea is that

- $\varphi((A+B)^k)$ can be expanded into terms of the form $\varphi(A^{k_1} B^{k_2} A^{k_3} B^{k_4} \dots)$
- such terms can in turn be expanded into lower order terms, using free independence. (see e.g. lemma 2)

#

Terminology:

The distribution of $A+B$ is called the additive free convolution of the distributions of A and B.

Question: In classical probability, the Fourier transform possesses the nice property that if μ and ν are two distributions, then $\phi_{\mu * \nu}(t) = \phi_\mu(t) \cdot \phi_\nu(t)$. Is there such a function for the free additive convolution?

Answer: yes; the so-called R-transform.

Defs:

- For $z \in \mathbb{C} \setminus \mathbb{R}$, we define the Stieltjes transform of A: (satisfying $A = A^T$)

$$g_A(z) := \varphi\left((A - zI)^{-1}\right)$$

- The R-transform of A is then defined as:

$$R_A(z) := g_A^{-1}(-z) - \frac{1}{z} \quad \text{(well defined } \forall z \in \mathbb{C}\text{, actually)}$$

(= inverse function, $\neq \frac{1}{g_A(z)}$)

Proposition (to be proven next time)

If A and B are freely independent, then $R_{A+B}(z) = R_A(z) + R_B(z)$.

(so to be precise, $R_A(z)$ plays the role of $\log \phi_\mu(t)$ here)

Example

Catalan numbers

- If B is distributed according to the semi-circle distribution (i.e. $m_k = \varphi(B^k) = \begin{cases} r_\ell & \text{if } k = 2\ell \\ 0 & \text{if } k = 2\ell+1 \end{cases}$) then we know that its Stieltjes transform satisfies the equation:

$$g_B(z)^2 + z\, g_B(z) + 1 = 0$$

i.e. $z = -g_B(z) - \frac{1}{g_A(z)}$

i.e. $g_B^{-1}(z) = -z - \frac{1}{z}$

× i.e. $R_B(z) = g_B^{-1}(-z) - \frac{1}{z} = z + \frac{1}{z} - \frac{1}{z} = z$

- So in this case, $R_{A+B}(z) = R_A(z) + z$

i.e. $g_{A+B}^{-1}(-z) - \frac{1}{z} = z + g_A^{-1}(-z) - \frac{1}{z}$

$-z \to g_{A+B}(z)$

i.e. $z = -g_{A+B}(z) + g_A^{-1}(g_{A+B}(z))$

i.e. $g_A(z + g_{A+B}(z)) = g_{A+B}(z)$

which was the composition rule found (in the limit $n \to \infty$) for the addition of a random matrix A and a GOE matrix $\mathbf{B}$, assumed to be independent.

Connection with random matrices

The connection follows from the following important proposition of Voiculescu:

Proposition

• Let $A^{(n)}$ and $B^{(n)}$ be two $\underbrace{\text{independent random matrices}}_{}$ (nxn real symmetric)
such that $F_n^A(t) \xrightarrow[n\to\infty]{} F^A(t)$ a.s., $F_n^B(t) \xrightarrow[n\to\infty]{} F^B(t)$ a.s.,
× with corresponding R-transforms $R_A(z)$ and $R_B(z)$.

• Let moreover $V^{(n)}$ be a nxn orthogonal matrix distributed according to the Haar distribution and independent of both $A^{(n)}$ and $B^{(n)}$.

• Then $A^{(n)}$ and $\tilde{B}^{(n)} = V^{(n)} B^{(n)} (V^{(n)})^T$ are asymptotically freely independent, i.e. $\forall m$,

$$\lim_{n\to\infty} \frac{1}{n} \mathbb{E}\left(\mathrm{Tr}\left(P_1(A^{(n)}) Q_1(\tilde{B}^{(n)}) \dots P_m(A^{(n)}) Q_m(\tilde{B}^{(n)})\right)\right) = 0$$

whenever $P_1 .. P_m$, $Q_1 .. Q_m$ are polynomials such that (*)

$$\lim_{n\to\infty} \frac{1}{n} \mathbb{E}\left(\mathrm{Tr}\, P_j(A^{(n)})\right) = \lim_{n\to\infty} \frac{1}{n} \mathbb{E}\left(\mathrm{Tr}\, Q_j(\tilde{B}^{(n)})\right) = 0 \quad \forall 1 \le j \le m$$

× • As a consequence, the limiting eigenvalue distribution of $A^{(n)} + \tilde{B}^{(n)} = A^{(n)} + V^{(n)} B^{(n)} (V^{(n)})^T$ is the distribution whose R-transform is given by $R_A(z) + R_B(z)$.

(*) here also, $P_1(A^{(n)}) = I$ or $Q_m(\tilde{B}^{(n)}) = I$ are also valid

Remarks

- The above proposition therefore also applies to $A^{(n)} + H^{(n)}$, where $A^{(n)}$ and $H^{(n)}$ are independent and $H^{(n)}$ is from the GOE (since $H^{(n)} = V^{(n)} \Lambda^{(n)} (V^{(n)})^T$, with $V^{(n)}$ distributed according to the Haar distribution, and $A^{(n)}$ is independent of both $\Lambda^{(n)}$ and $V^{(n)}$, which are in turn mutually independent (see Lecture 2)).

- We have found here a composition rule for random matrices whose eigenvectors have no relation whatsoever between them. This is coherent with the fact seen last time that whenever two matrices share the same set of eigenvectors, then no general composition rule exists.