Lecture 27: June 3, 2011

1. Mixed Matrix Moments

This final lecture is devoted to calculating the joint moments of a GUE_n matrix X_n and any *deterministic* (i.e. constant) matrix $D_n \in (\mathcal{M}_n, \varphi_n)$. In other words, we want to calculate the limits of all moments of the form

$$\varphi_n(X_n D_n^{q_1} X_n D_n^{q_2} \cdots X_n D_n^{q_k}) \tag{Moment}$$

for any $k \geq 1$ and any non-negative integers $q_1, \ldots, q_k \in \mathbb{N}$. (Since we allow some or all of the q_i to equal 0, and since φ_n is tracial, any joint moment can be written in this form.) To simplify notation, denote

$$[D_n^{q_i}]_{jk} = d_{jk}^{(i)}.$$

Of course, we can calculate then that $d_{jk}^{(i)} = \sum_{1\leq j_1,\ldots,j_{q_i-1}\leq n} D_{jj_1} D_{j_1j_2} \cdots D_{j_{q_i-1}k}$, and so the $d_{jk}^{(i)}$ are all determined by the D_{jk}. However, as we will soon see, it is more convenient to treat the $d^{(i)}$ as separate quantities. Hence, suppressing the n in the notation as in the previous lecture, we have the general mixed moment of Equation Moment is equal to

$$\frac{1}{n}\mathbb{E}\,\mathrm{Tr}\,(XD^{q_1}XD^{q_2}\cdots XD^{q_k}) = \frac{1}{n}\sum_{\substack{1\leq i_1,\ldots,i_k\leq n\\ 1\leq j_1,\ldots,j_k\leq n}} \mathbb{E}(X_{i_1j_1} d_{j_1i_2}^{(1)} X_{i_2j_2} d_{j_2i_3}^{(2)} \cdots X_{i_kj_k} d_{j_ki_1}^{(k)}).$$

We may simply factor out all the constant $d_{jk}^{(i)}$ terms, to give a sum of terms of the form

$$d_{j_1i_2}^{(1)} d_{j_2i_3}^{(2)} \cdots d_{j_ki_1}^{(k)} \mathbb{E}(X_{i_1j_1}X_{i_2j_2}\cdots X_{i_kj_k}).$$

The expectation can be calculated using Wick's formula, since the entries form a Gaussian family; we have

$$\mathbb{E}(X_{i_1j_1}X_{i_2j_2}\cdots X_{i_kj_k}) = \sum_{\pi\in\mathscr{P}_2(k)} \prod_{\{a,b\}\in\pi} \mathbb{E}(X_{i_aj_a}X_{i_bj_b}) = \sum_{\pi\in\mathscr{P}_2(k)} \frac{1}{n^{k/2}} \prod_{\{a,b\}\in\pi} \delta_{i_aj_b}\delta_{i_bj_a},$$

where the last equality follows from Equation (cov) in Lecture 25. Hence, we may express (Moment) as

$$n^{-k/2-1} \sum_{\pi\in\mathscr{P}_2(k)} \sum_{\substack{1\leq i_1,\ldots,i_k\leq n\\ 1\leq j_1,\ldots,j_k\leq n}} \prod_{\{a,b\}\in\pi} \delta_{i_aj_b}\delta_{i_bj_a} d_{j_1i_2}^{(1)} d_{j_2i_3}^{(2)} \cdots d_{j_ki_1}^{(k)}.$$

As usual, it is now convenient to identify $\pi \in \mathscr{P}_2(k)$ with the partition P_π. For fixed $i_1, \ldots, i_k, j_1, \ldots, j_k$, the product then simply yields that the only non-zero terms are those with $i_a = j_{P_\pi(a)}$ for each $a \in [k]$; for convenience, we write this as $i_a = j_{\pi(a)}$. So, utilizing the clockwise one-step permutation γ, we can rewrite (Moment) as

$$n^{-k/2-1} \sum_{\pi\in\mathscr{P}_2(k)} \sum_{1\leq j_1,\ldots,j_k\leq n} d_{j_1j_{\pi\gamma(1)}}^{(1)} d_{j_2j_{\pi\gamma(2)}}^{(2)} \cdots d_{j_kj_{\pi\gamma(k)}}^{(k)}. \tag{1.1}$$

In order to recognize this sum as a familiar object, let us reorder the terms in the internal product according to the cycles of $\pi\gamma$. For example, suppose $k = 6$ and take

$\pi = (1,3)(2,5)(4,6)$. Then $\pi\gamma = (1,5,4,2)(3,6)$. So we write the 6-term product as

$$\begin{aligned} d^{(1)}_{j_1 j_{\pi\gamma(1)}} d^{(2)}_{j_2 j_{\pi\gamma(2)}} \cdots d^{(6)}_{j_6 j_{\pi\gamma(6)}} &= d^{(1)}_{j_1 j_5} d^{(2)}_{j_2 j_1} d^{(3)}_{j_3 j_6} d^{(4)}_{j_4 j_2} d^{(5)}_{j_5 j_4} d^{(6)}_{j_6 j_3} \\ &= d^{(1)}_{j_1 j_5} d^{(5)}_{j_5 j_4} d^{(4)}_{j_4 j_2} d^{(2)}_{j_2 j_1} \cdot d^{(3)}_{j_3 j_6} d^{(6)}_{j_6 j_3}. \end{aligned}$$

Now summing over the 6 indices, we have the internal sum is equal to

$$\sum_{1 \le j_1, \ldots, j_6 \le n} d^{(1)}_{j_1 j_5} d^{(5)}_{j_5 j_4} d^{(4)}_{j_4 j_2} d^{(2)}_{j_2 j_1} \cdot d^{(3)}_{j_3 j_6} d^{(6)}_{j_6 j_3} = \sum_{1 \le j_1, j_5, j_4, j_2 \le n} d^{(1)}_{j_1 j_5} d^{(5)}_{j_5 j_4} d^{(4)}_{j_4 j_2} d^{(2)}_{j_2 j_1} \cdot \sum_{1 \le j_3, j_6 \le n} d^{(3)}_{j_3 j_6} d^{(6)}_{j_6 j_3}.$$

Each of these is a trace: in this case, we get

$$\mathrm{Tr}\,(D^{(1)} D^{(5)} D^{(4)} D^{(2)}) \cdot \mathrm{Tr}\,(D^{(3)} D^{(6)}),$$

where we write $D^{(j)} = D^{q_j}$. In this way, all the terms will be products of traces. Let us further express this in terms of *normalized traces*: let $\mathrm{tr} = \frac{1}{n}\,\mathrm{Tr}$. Then in the above example, we get

$$n^{\#(\pi\gamma)}\,\mathrm{tr}\,(D^{(1)} D^{(5)} D^{(4)} D^{(2)}) \cdot \mathrm{tr}\,(D^{(3)} D^{(6)}).$$

Following identical reasoning, we get the following in general.

Lemma 1. *Let $D^{(1)}, \ldots, D^{(k)}$ be $n \times n$ matrices, with $d^{(i)}_{jk}$ denoting the jk-entry of $D^{(j)}$. For any $\sigma \in S_k$,*

$$\sum_{1 \le j_1, \ldots, j_k \le n} d^{(1)}_{j_1 j_{\sigma(1)}} d^{(2)}_{j_1 j_{\sigma(2)}} \cdots d^{(k)}_{j_k j_{\sigma(k)}} = n^{\#\sigma}\,\mathrm{tr}_\sigma(D^{(1)}, \ldots, D^{(k)}).$$

The notation in Lemma 1 is in terms of constructions we're familiar with. The symbol $\#\sigma$ denotes the number of cycles in σ. The symbol tr_σ is the multiplicative extension of the multi-linear functional $(D^{(1)}, \ldots, D^{(k)}) \mapsto \mathrm{tr}(D^{(1)} \cdots D^{(k)})$ over the *cycles* of the permutation σ. That is, if $B = (i_1, i_2, \ldots, i_s)$ is a cycle, denote

$$\mathrm{tr}(B)(D^{(1)}, \ldots, D^{(k)}) = \mathrm{tr}(D^{(i_1)} D^{(i_2)} \cdots D^{(i_s)}).$$

Then if $\sigma = B_1 \cdot B_2 \cdots B_r$ is a product of independent cycles (where we *include* cycles of length 1), $\mathrm{tr}_\sigma = \mathrm{tr}(B_1) \cdots \mathrm{tr}(B_r)$. This is *precisely* the multiplicative extension of the linear functionals $\varphi_k \colon (a_1, \ldots, a_k) \mapsto \varphi(a_1 \cdots a_k)$ we defined for an arbitrary NCPS $(\mathscr{A}, \varphi)$, in the special case $(\mathscr{A}, \varphi) = (M_n(\mathbb{C}), \mathrm{tr})$. Previously, we only extended the multiplicative structure over $\sigma \in \mathscr{P}(k)$; by allowing the order of the variables within each block to be non-increasing, the same definition extends to all $\sigma \in S_k$.

Hence, combining Lemma 1 with Equation 1.1, we arrive at a useful expression for the mixed moments of a GUE_n with any constant matrix D.

Proposition 2. *The mixed moments of the GUE_n matrix $X = X_n$ with any constant matrix $D = D_n$ are given by*

$$\varphi_n(XD^{q_1}XD^{q_2} \cdots XD^{q_k}) = \sum_{\pi \in \mathscr{P}_2(k)} n^{\#(\pi\gamma) - k/2 - 1}\,\mathrm{tr}_{\pi\gamma}(D^{q_1}, \ldots, D^{q_k}).$$

Proposition 2 holds for any fixed finite n. Note that it has the feeling of a genus expansion: the exponent of $n^{\#(\pi\gamma) - k/2 - 1}$ is the same one that appears there. (This is because $\#(\pi\gamma) = \#(\gamma\pi)$. In general, for any permutations σ, α, the permutation $\alpha\sigma\alpha^{-1}$ has the same cycle structure as σ; indeed, as one can easily check, $(i_1, i_2, \ldots, i_s)$ is a cycle of σ if and only if $(\alpha(i_1), \alpha(i_2), \ldots, \alpha(i_s))$ is a cycle of $\alpha\sigma\alpha^{-1}$. In particular, this means $\#(\alpha\sigma\alpha^{-1}) = \#\sigma$. Thus, for any permutations α, β, we have $\#(\beta\alpha) = \#(\alpha \cdot \beta\alpha \cdot \alpha^{-1}) = \#(\alpha\beta)$.)

The idea now is to let $n \to \infty$. As usual, the only surviving terms will be those for which $\#(\pi\gamma) - k/2 - 1 = 0$, which means $\pi \in NC_2(k)$. But remember that the matrix $D = D_n$ also depends on n, so we need to know that the terms $\mathrm{tr}_{\pi\gamma}(D_n^{q_1}, \ldots, D_n^{q_k})$ have limits as well.

2. Asymptotic Moments

The assumption we need to make, of course, is that the sequence $D_n \in (\mathcal{M}_n, \varphi_n)$ has a limit distribution. Note: since D_n is deterministic, $\varphi_n = \frac{1}{n}\mathbb{E}\,\mathrm{Tr} = \frac{1}{n}\,\mathrm{Tr} = \mathrm{tr}$ on the algebra generated by D_n. Let us restrict to the case $D_n = D_n^*$. (This is not necessary to prove the convergence results, but it is the interesting case.) Then the assumption we must make is that there is a NCPS $(\mathscr{A}, \varphi)$ containing a self-adjoint element $d \in \mathscr{A}$ such that

$$D_n \xrightarrow{\mathscr{D}} d, \quad \text{as } n \to \infty.$$

In other words, for any $p \in \mathbb{N}$, $\mathrm{tr}(D_n^p) = \varphi_n(D_n^p) \to \varphi(d^p)$ as $n \to \infty$. Supposing this is the case, let $k \in \mathbb{N}$, let $\sigma \in S_k$ be any permutation, and let $q_1, \ldots, q_k \in \mathbb{N}$. Consider the sequence of numbers $\mathrm{tr}_\sigma(D_n^{q_1}, \ldots, D_n^{q_k})$. If σ has cycles $B_1, \ldots, B_r$, then there are non-negative integers $p_1, \ldots, p_r$ with

$$\mathrm{tr}_\sigma(D_n^{q_1}, \ldots, D_n^{q_k}) = \mathrm{tr}(D_n^{p_1}) \cdots \mathrm{tr}(D_n^{p_r});$$

indeed, by definition $\mathrm{tr}_\sigma = \mathrm{tr}(B_1) \cdots \mathrm{tr}(B_r)$, and if $B = (i_1, \ldots, i_s)$ then

$$\mathrm{tr}(B)(D_n^{q_1}, \ldots, D_n^{q_k}) = \mathrm{tr}(D_n^{q_{i_1}} \cdots D_n^{q_{i_s}}) = \mathrm{tr}(D_n^{q_{i_1} + \cdots + q_{i_s}}).$$

Hence, the assumption that $D_n \xrightarrow{\mathscr{D}} d$ implies that, for any $\sigma \in S_k$,

$$\mathrm{tr}_\sigma(D_n^{q_1}, \ldots, D_n^{q_k}) = \mathrm{tr}(D_n^{p_1}) \cdots \mathrm{tr}(D_n^{p_r}) \longrightarrow \varphi(d^{p_1}) \cdots \varphi(d^{p_r}) = \varphi_\sigma(d^{q_1}, \ldots, d^{q_k}) \quad \text{as } n \to \infty.$$

Combining this with Proposition 2 and the observation from the genus expansion that

$$\text{for any } \pi \in \mathscr{P}_2(k), \lim_{n\to\infty} n^{\#(\pi\gamma) - k/2 - 1} = \mathbb{1}_{NC_2(k)}(\pi),$$

we have the following.

Proposition 3. *If X_n is a GUE_n and D_n is a sequence of self-adjoint constant matrices in $(\mathcal{M}_n, \varphi_n)$ with $D_n \xrightarrow{\mathscr{D}} d \in (\mathscr{A}, \varphi)$, then for any $q_1, \ldots, q_k \in \mathbb{N}$*

$$\lim_{n\to\infty} \varphi_n(X_n D_n^{q_1} X_n D_n^{q_2} \cdots X_n D_n^{q_k}) = \sum_{\pi \in NC_2(k)} \varphi_{\pi\gamma}(d^{q_1}, d^{q_2}, \ldots, d^{q_k}).$$

Now recall from Lecture 26: if $a_1, b_1, \ldots, a_k, b_k$ are elements of a NCPS $(\mathscr{A}, \varphi)$ with $\{a_1, \ldots, a_k\}$ and $\{b_1, \ldots, b_k\}$ freely independent, the mixed moments can be calculated as

$$\varphi(a_1 b_1 \cdots a_k b_k) = \sum_{\pi \in NC(k)} \kappa_\pi(a_1, \ldots, a_k) \varphi_{K(\pi)}(b_1, \ldots, b_k).$$

Let us specialize to the case that $a_1 = \ldots = a_k = s$ is a semicircular random variable. Then $\kappa_\pi(a_1, \ldots, a_k) = \kappa_\pi(s, \ldots, s)$ is 0 unless all the blocks of π have size 2, in which case it is equal to 1. In other words,

$$\varphi(sb_1sb_2\cdots sb_k) = \sum_{\pi\in NC_2(k)} \varphi_{K(\pi)}(b_1,\ldots,b_k).$$

On the other hand, also proved in Lecture 26, we saw that the Kreweras complement $K(\pi)$ can be identified as a permutation by $P_{K(\pi)} = P_\pi^{-1}\gamma$. If π is a pairing, $P_\pi = P_\pi^{-1}$, and so (identifying π with P_π as we have been doing) this gives

$$\varphi(sb_1sb_2\cdots sb_k) = \sum_{\pi\in NC_2(k)} \varphi_{\pi\gamma}(b_1,\ldots,b_k). \tag{2.1}$$

So, finally, combining with Proposition 3, we have arrived at our main theorem.

Theorem 4. *If X_n is a GUE_n and D_n is a sequence of self-adjoint constant matrices in $(\mathcal{M}_n,\varphi_n)$ with $D_n \xrightarrow{\mathscr{D}} d \in (\mathscr{A},\varphi)$, then there is a semicircular variable s (in a potentially larger NCPS)* free *from d so that*

$$(X_n, D_n) \xrightarrow{\mathscr{D}} (s,d).$$

In particular, X_n and D_n are asymptotically free.

Proof. The statement is that, if Q is any non-commutative polynomial in 2-variables, then $\varphi_n(Q(X_n,D_n)) \to \varphi(Q(s,d))$ as $n\to\infty$. By linearity it suffices to prove this for monomials Q, and (as explained above) by allowing powers $q_1,\ldots,q_k$ to take on any non-negative values (including 0), it suffices to prove this for $Q(x,y) = xy^{q_1}xy^{q_2}\cdots xy^{q_k}$ for $k\geq 1$. (The case of monomials starting with a y need not be considered separately since φ_n is tracial, so $\varphi_n(D_n^{q_0}X_nD_n^{q_1}X_n\cdots X_nD_n^{q_k}) = \varphi_n(X_nD_n^{q_1}X_n\cdots X_nD_n^{q_k+q_0})$.) As Proposition 3 shows,

$$\lim_{n\to\infty}\varphi_n(Q(X_n,D_n)) = \lim_{n\to\infty}\varphi_n(X_nD_n^{q_1}\cdots X_nD_n^{q_k}) = \sum_{\pi\in NC_2(k)}\varphi_{\pi\gamma}(d^{q_1},\ldots,d^{q_k}).$$

Since s is free from d, s is free from $\{d^{q_1},\ldots,d^{q_k}\}$. Equation 2.1 therefore shows that

$$\sum_{\pi\in NC_2(k)}\varphi_{\pi\gamma}(d^{q_1},\ldots,d^{q_k}) = \varphi(sd^{q_1}\cdots sd^{q_k}) = \varphi(Q(s,d)).$$

This proves the result. □

3. WHY THIS IS A BIG DEAL

The distribution of a GUE_n $X_n \in (\mathcal{M}_n,\varphi_n)$ can be interpreted as a measure on $\mathbb{R}$, as we have seen: it is the (averaged) empirical eigenvalue distribution μ_{X_n} of the random matrix. Similarly, if D_n is any self-adjoint constant matrix, its distribution in $(\mathcal{M}_n,\varphi_n)$ is the (fixed) eigenvalue distribution μ_{D_n}:

$$\mu_{D_n} = \frac{1}{n}\sum_{\lambda\in\mathrm{Eig}(D_n)}\delta_\lambda.$$

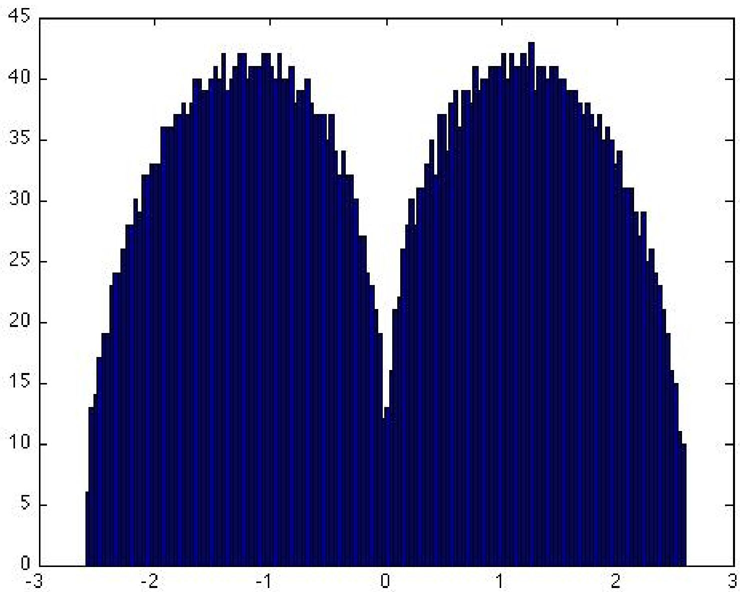
We know that $\mu_{X_n} \to S(0,1)$ (the standard semicircle law), while $\mu_{D_n}\to\mu$ for some measure μ (determined by its moments, the moments $\varphi(d^m)$). In this light, Theorem 4 tells us that

$$\mu_{D_n+X_n} \to \mu \boxplus S(0,1) \quad \text{as } n\to\infty.$$

In other words, if we start with a deterministic matrix D_n, and then perturb it by adding the random Gaussian matrix X_n, the eigenvalues of the "noisy" matrix $D_n + X_n$ are approximated by a predictable law! Let's consider a specific example.

Example. Suppose n is even, and let D_n be a diagonal matrix with $[D_n]_{jj} = 1$ if $1 \le j \le n/2$ and $[D_n]_{jj} = -1$ if $n/2 < j \le n$. Then D_n has only two distinct eigenvalues, ± 1, and it is easy to see that $\mu_{D_n} = \frac{1}{2}(\delta_1 + \delta_{-1})$ is the symmetric Bernoulli law μ.

What happens to the eigenvalues when we add a GUE_n to D_n? The following plot is the actual histogram of eigenvalues of an instance of $D_n + X_n$ with $n = 5000$, produced using MATLAB.



The preceding discussion suggests that this law should be well-approximated by $\mu \boxplus S(0,1)$. This is something we can compute analytically. As calculated in Lecture 17, the $\mathscr{R}$-transforms of μ and $S(0,1)$ are

$$\mathscr{R}_\mu(z) = \frac{-1 + \sqrt{1 + 4z^2}}{2z} \qquad \mathscr{R}_{S(0,1)}(z) = z.$$

Thus, $\mathscr{R}_{\mu \boxplus S(0,1)}$ is the sum of these two for all sufficiently small $|z|$. Since we also know that $G_\nu(\mathscr{R}_\nu(z) + 1/z) = z$, we can compute the Stieltjes transform of $\mu \boxplus S(0,1)$ by solving

$$G_{\mu \boxplus S(0,1)}\left(z + \frac{-1 + \sqrt{1 + 4z^2}}{2z} + \frac{1}{z} \right) = z.$$

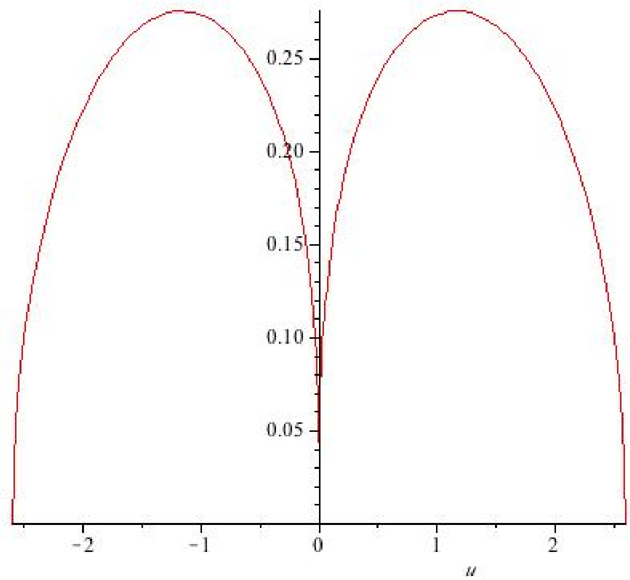
Set $u = z + \frac{-1+\sqrt{1+4z^2}}{2z} + \frac{1}{z}$; so $2uz = 2z^2 + 1 + \sqrt{1 + 4z^2}$. Solving for z in terms of u involves a nominally quartic equation, but a z factors out and the equation becomes $z(z-u)^2 = u$; that is

$$z^3 - 2uz^2 + u^2 z - u = 0.$$

One can use the cubic formula to solve for $z = G_{\mu\boxplus S(0,1)}(u)$ from here. It is somewhat tricky, since there are three roots and the correct one must be identified (from a complicated formula). However, upon identifying regions where the imaginary part is non-zero for $u \in \mathbb{R}$, one can then calculate using the Stieltjes inversion formula that

$$(\mu \boxplus S(0,1))(du) = \frac{1}{\pi}\left(\frac{1}{6}f(u) - \frac{2}{3}\frac{u^2}{f(u)}\right)\mathbb{1}_{|u|\le\frac{3}{2}\sqrt{3}}\,du,$$

where $f(t) = (108|u| - 8|u|^3 + 12|u|\sqrt{81-12|u|^2})^{1/3}$. This is not a simple formula to analyze, but graphing this density demonstrates Theorem 4 magnificently:



Exercise. For $\lambda > 0$, let $\mu_\lambda = \frac{1}{2}(\delta_\lambda + \delta_{-\lambda})$. Let $S(0,t)$ denote the semicircle law of variance t. Calculate the density of the measure $\mu_\lambda \boxplus S(0,t)$, and graph it. Then use MATLAB to calculate the eigenvalues of $D_n + tX_n$ where X_n is a GUE_n and D_n is diagonal with $n/2$ entries equal to λ and the other $n/2$ entries equal to $-\lambda$.

While this looks like magic, it still may not be clear why it is a BIG DEAL. The reason is this. Suppose that a radio signal is sent out. It is detected on an array of n antennas, which sample at regular intervals (on the order of milliseconds). If m samples are taken in this "time-series", the resulting signal is an $n \times m$ matrix (of signal strengths). Of course, what is received is not exactly what was sent out: there is "noise in the channel". The way that engineers model this noise is as follows: the signal is a fixed matrix $D_{n,m}$, and what is received is $Y_{n,m} = D_{n,m} + X_{n,m}$ where $X_{n,m}$ is a random $n \times m$ matrix. The most common model to use for $X_{n,m}$ is to assume that its entries are independent normals.

The situation of Theorem 4 is a little more restrictive: it requires $n = m$, and the matrix $X_n = X_{n,n}$ has independent normals above the main diagonal but is self-adjoint; similarly, $D_{n,n} = D_n$ is assumed self-adjoint. These are unphysical assumptions, so we are not quite in a position to help the engineers. But the above exposition is just the beginning. One can consider the general $n \times m$ case, and then look at the matrix $Y_{n,m}Y_{n,m}^*$. Studying the eigenvalues of this matrix is known as PCA: Principal Component Analysis. It turns out that the tools of free probability also provide insight into this situation: one can predict exactly what the distribution of eigenvalues of $Y_{n,m}$ should look like if n, m are large; then, by comparing to the actual measured signal, one can "deconvolve" to recover the original message. Algorithms based on this kind of idea are currently being developed in signal processing labs (at MIT and Princeton, for example), with the likely output being *dramatic* improvements in the performance of algorithms that, for example, keep your cell-phone connected to a tower.

4. Extensions

Theorem 4 has an immediate extension. With exactly the same proof (with slightly more complicated notation), we have:

Theorem 5. *Let $\{X_n^p\}_{p\in I}$ be a family of independent GUE_n matrices. Let $\{D_n^q\}_{q\in J} \in (\mathcal{M}_n, \varphi_n)$ be a collection of constant matrices that possesses a limit distribution: $\{D_n^q\}_{q\in J} \xrightarrow{\mathscr{D}} \{d_q\}_{q\in J}$ for some variables d_q in a NCPS. Then*

$$\{X_n^p, D_n^q\}_{p\in I, q\in J} \xrightarrow{\mathscr{D}} \{s_p, d_q\}_{p\in I, q\in J}$$

where $\{s_p\}_{p\in I}$ is a semicircular system freely independent from $\{d_q\}_{q\in J}$. In particular, $\{X_n^p\}_{p\in I}$ and $\{D_n^q\}_{q\in J}$ are asymptotically free.

This theorem still does not really get us beyond the world of Gaussians completely, though. One is left wondering if asymptotic freeness can only really occur in the presence of (approximate) semicircular systems. The answer is no: it abounds quite "freely".

The key observation for what really makes for asymptotic freeness comes from the diagonalization of a GUE_n matrix. This is a topic for my Fall 2011 Math 247A class: Random Matrices. The result is as follows.

Proposition 6. *Let X_n be a GUE_n. Let $X_n = U_n^*\Lambda_n U_n$ be the unitary diagonalization. The joint distribution of the random matrix U_n is the Haar measure on the unitary group $U(n)$.*

In the proposition Λ_n is diagonal with the eigenvalues of X_n on the diagonal, and U_n is the eigenvector matrix, which is in the unitary group $U(n)$. This diagonalization is uniquely defined up to the order of the eigenvalues; the result does not depend on which order is chosen, but to be definite, we canonically list them in increasing order.

The Haar measure is not easy to describe in terms of a density for the joint distribution, but it is a natural group-theoretic object. The group $U(n)$ is a compact topological group (in fact a Lie group). On any such group, there is a *unique* (regular) Borel probability measure that is invariant under the (left) action of the group. That is: if G is a compact topological group, there is a unique regular probability measure μ_G on the Borel σ-field of G with the property that, for any Borel set S and any $x \in G$, $\mu_G(xS) = \mu_G(S)$. For example, take $G = U(1)$, which is the circle group. Then $\mu_{U(1)}(d\theta) = \frac{1}{2\pi}\, d\theta$; this probability measure is invariant under rotations.

The fact that the eigenvector matrix of a GUE_n has the Haar distribution is a reflection of the remarkable rotational-invariance properties of Gaussian distributions. In general, a U_n distributed according to $\mu_{U(n)}$ is called a **Haar unitary** or **Haar unitary random matrix** (if we need to distinguish it from Haar unitary operators in a NCPS). In the next course, we will show that if U_n is a sequence of $n \times n$ Haar unitary random matrices in $(\mathcal{M}_n, \varphi_n)$, then $U_n \xrightarrow{\mathscr{D}} u$ where u is a Haar unitary (i.e. u is unitary and $\mu_u = \mu_{U(1)}$, the uniform measure on the unit circle in the complex plane).

The next theorem shows that it is this property of GUE_n eigenvectors that really makes them such ideal candidates for asymptotic freeness.

Theorem 7. *Let A_n, B_n be constant matrices in $(\mathcal{M}_n, \varphi_n)$ that (separately) have limit distributions: there exist a, b in NCPS's such that $A_n \xrightarrow{\mathscr{D}} a$ and $B_n \xrightarrow{\mathscr{D}} b$. Let U_n be Haar unitary random matrices. Then there is an NCPS containing* free *elements a', b' such that $\mu_{a'} = \mu_a$, $\mu_{b'} = \mu_b$, and*

$$(A_n, U_n^* B_n U_n) \xrightarrow{\mathscr{D}} (a', b').$$

In particular, A_n and $U_n^* B_n U_n$ are asymptotically free.

Remark. If u is any unitary in any tracial NCPS and a is any random variable, it is clear that $\mu_a = \mu_{u^*au}$ (just compute the moments, and observe that all u's cancel with u^*'s internally). In particular, $U_n^* B_n U_n$ has the same non-commutative distribution as B_n (meaning its eigenvalues are unchanged).

This result finally *truly* tells us what (asymptotic) freeness means. To say that two matrices are asymptotically free is to say that the eigenvectors of one are *uniformly randomly rotated* relative to the other. This theorem also answers the question (posed in Lecture 21) of how we can generate "free coin tosses". First, let A_n be a matrix whose eigenvalue distribution is the Bernoulli law (for example, as above, take A_n diagonal with $1/2$ its diagonal entries $= 1$ and the other half $= -1$). Then if U_n is a sequence of Haar unitaries, we find that the two matrices $A_n, U_n^* A_n U_n$ are asymptotically free, each having the Bernoulli law as their distribution. (Thus, the empirical eigenvalue distribution of $A_n + U_n^* A_n U_n$ is close to the arcsine law, following Lecture 21.)

Theorem 7 can be proved using combinatorial methods following those we developed this term. If we had an additional 10 lectures, we could prove this theorem; for now, its proof will have to remain a mystery.

Finally, what about truly *random* matrices and asymptotic freeness? In Theorems 4 and 7, the matrices A_n, B_n, D_n are all constant matrices. One may naturally wonder if they could be replaced with random matrices? The answer is yes, but only if the notion of convergence is correct. Let's take a look at two different ways to talk about convergence.

Definition 8. *Let $A_n \in (\mathcal{M}_n, \varphi_n)$ be a sequence of random matrices. When we say $A_n \to a \in (\mathscr{A}, \varphi)$, we might mean two different things. The literal definition is that $\varphi_n(P(A_n)) \to \varphi(P(a))$ for all polynomials P. An alternative would be to consider* almost sure *convergence in the* smaller *NCPS $(M_n(\mathbb{C}), \mathrm{tr}_n)$. That is to say: $A_n \to a$ if and only if for every polynomial and almost every ω, $\mathrm{tr}_n(P(A_n(\omega))) \to \varphi(P(a))$. Call this* almost sure convergence in distribution.

Almost sure convergence in distribution is (morally) stronger than convergence in distribution. Theorems 4 and 7 are true for *random* A_n, B_n, D_n (with appropriate *independence* assumptions) *provided* their convergence in distribution is in the almost-sure sense (and then the conclusion of the theorem is again almost-sure convergence). For example,

Theorem 9. *Let $A_n, B_n \in (\mathcal{M}_n, \varphi_n)$ be random matrices that (separately) have a.s. limit distributions: there exist a, b in NCPS's such that $A_n \xrightarrow{\mathscr{D}} a$ a.s. and $B_n \xrightarrow{\mathscr{D}} b$ a.s. Let U_n be Haar unitary* *random matrices* independent from A_n. *Then there is an NCPS containing* free *elements* a', b' *such that* $\mu_{a'} = \mu_a$, $\mu_{b'} = \mu_b$, *and*

$$(A_n, U_n^* B_n U_n) \xrightarrow{\mathscr{D}} (a', b') \ \ a.s.$$

In particular, A_n *and* $U_n^* B_n U_n$ *are a.s. asymptotically free.*

It is important to note that independence alone is not enough to give asymptotic freeness: in general, even if A_n and B_n are independent and have a.s. limit distributions, they may not be asymptotically free. Moreover, the more straightforward notion of convergence in distribution in $(\mathcal{M}_n, \varphi_n)$ (in terms of expectations) does *not* yield asymptotic freeness with random A_n, B_n, D_n. Studying the proof of Theorem 4, we can see why. The key to that proof is Proposition 2,

$$\varphi_n(XD^{q_1}XD^{q_2}\cdots XD^{q_k}) = \sum_{\pi\in\mathscr{P}_2(k)} n^{\#(\pi\gamma)-k/2-1}\,\mathrm{tr}_{\pi\gamma}(D^{q_1},\ldots,D^{q_k})$$

when D is constant. If D is random, the result would be

$$\varphi_n(XD^{q_1}XD^{q_2}\cdots XD^{q_k}) = \sum_{\pi\in\mathscr{P}_2(k)} n^{\#(\pi\gamma)-k/2-1}\mathbb{E}\left(\,\mathrm{tr}_{\pi\gamma}(D^{q_1},\ldots,D^{q_k})\right).$$

The issue is that

$$\mathbb{E}\left(\,\mathrm{tr}_{\pi\gamma}(D^{q_1},\ldots,D^{q_k})\right) \neq (\varphi_n)_{\pi\gamma}(D^{q_1},\ldots,D^{q_k}).$$

Indeed, consider the example above where $\pi\gamma = (1,5,4,2)(3,6)$. In this case,

$$\mathbb{E}\left(\,\mathrm{tr}_{\pi\gamma}(D^{q_1},\ldots,D^{q_k})\right) = \mathbb{E}\left(\,\mathrm{tr}(D^{q_1}D^{q_5}D^{q_4}D^{q_2})\,\mathrm{tr}(D^{q_3}D^{q_6})\right)$$

while

$$(\varphi_n)_{\pi\gamma}(D^{q_1},\ldots,D^{q_k}) = \mathbb{E}\,\mathrm{tr}(D^{q_1}D^{q_5}D^{q_4}D^{q_2})\cdot\mathbb{E}\,\mathrm{tr}(D^{q_3}D^{q_6}).$$

In general, these are not equal. One might hope that the expected traces do *converge in expectation* to the product of the expected traces; such convergence issues are termed *fluctuations*. The fact is, only assuming that $\mathbb{E}\,\mathrm{tr}(D_n^q)$ converges for each q is not sufficient to prove this convergence: the fluctuations may still be large enough that "sub-dominant" terms contribute enough to make things blow up. Thus, almost sure convergence of the (random) moments $\mathrm{tr}(D_n^q)$ is required. As usual, the proof will involve showing that the associated variance tends to 0. We will discuss problems like this one in Math 247A: Random Matrices in Fall 2011.