Piotr Śniady

IMPAN lectures 2017/2018

Free probability and random matrices

Lecture 1. October 5, 2017
GUE random matrices

CHANGELOG

- the previous version of this file had the letters n and N mixed. :(

  The correct convention:
  N is for the SIZE of the random matrix.

- more consistent usage of Y for a GUE instead of X.

Empirical eigenvalues distribution.

suppose X is a $N \times N$ random matrix with

(random) eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_N \in \mathbb{C}$

$$\mu_X = \frac{1}{N} \sum_i \delta_{\lambda_i}$$

Counting measure on the set of the eigenvalues.

RANDOM probability measure on $\mathbb{C}$
gives full information about the eigenvalues of X.

TYPICAL GOAL: show some "law of large numbers" for μ_X as $N \to \infty$

Moments method:

$$\int_{\mathbb{C}} z^k \, d\mu_X = \underbrace{\frac{1}{N} \operatorname{Tr} X^k}_{\text{random variable!}}$$

random variable.

if X is not hermitian, this kind of information is NOT helpful.

if $X = X^*$ is hermitian, $\lambda_1, \dots, \lambda_N \in \mathbb{R}$ and there is some chance of success

$$\mathbb{E} \int_{\mathbb{C}} z^k \, d\mu_X = \int_{\mathbb{C}} z^k \underbrace{\mathbb{E}\mu_X}_{\substack{\text{non-random} \\ \text{probability} \\ \text{measure on } \mathbb{C}}} = \mathbb{E} \frac{1}{N} \operatorname{Tr} X^k$$

good for understanding the average eigenvalues distribution.

$$\operatorname{Var} \int_{\mathbb{C}} z^k \, d\mu_X = \operatorname{Var} \frac{1}{N} \operatorname{Tr} X^k$$

if $X = X^*$ is hermitian, this is a real-valued random variable.

Small variance = law of large numbers.

moments and convergence of measures

$\longrightarrow$ [MS] Section 2.1.

probability measures on $\mathbb{R}$

moments

convergence in moments

weak convergence of probability measures

measures determined by moments

Carleman's criterion

Hint: if $\sum \frac{1}{\sqrt[2n]{m_{2n}}} = +\infty$ then moment problem is determinate

convergence in moments vs weak convergence

convergence in moments is usually disdained by probabilists. For example: the limit might be not unique.

However, in the non-commutative setup we cannot really formulate "weak convergence of probability measures" and we have to stick to moments.

Magic of centered Gaussian distributions.

Claim. Suppose joint distribution of the family (X_α) of random variables is centered Gaussian. Then

$$\mathbb{E}\, X_{\alpha_1} X_{\alpha_2} \cdots X_{\alpha_n} = \sum_{\pi = \{\{\pi_{1,1}, \pi_{1,2}\}, \dots\}} \prod_i \underbrace{\mathbb{E}\, X_{\pi_{i,1}} X_{\pi_{i,2}}}_{\text{covariance}}$$

π is a pairing of $[n] := \{1, 2, \dots, n\}$

Exercise: prove it!

$\longrightarrow$ [MS] Section 1.4.

complex Gaussian distribution.

random variable Z has complex Gaussian distibution

if joint distribution of $X = \mathrm{Re}\, Z$ and $Y = \mathrm{Im}\, Z$ is

Gaussian.

our favorite example : $Z = X + iY$
with $X, Y \sim N(0, \delta^2)$ and independent.

$$\mathbb{E}\, Z = 0$$

$$\mathbb{E}\, Z^2 = \mathbb{E}\, X^2 - Y^2 + 2i\, XY = 0$$

$$\mathbb{E}\, \bar{Z}^2 = 0$$

$$\mathbb{E}\, Z\bar{Z} = \mathbb{E}\, |Z|^2 = \mathbb{E}\, X^2 + Y^2 = 2\delta^2.$$

if you work with complex-valued random variables, covariances involve random variables AND their complex conjugates.

afraid of complex random variables? always can translate everything to Re... and Im...

Ginibre ensemble.

X is a $N \times N$ matrix with entries (f_{ij})

$(\mathrm{Re}\, f_{ij}, \mathrm{Im}\, f_{ij})_{ij}$ – family of independent Gaussian random variables $\sim N(0, \delta^2)$

Exercise. Write a computer program in some fancy language which → generates a Ginibre random matrix, → calculates its (complex!) eigenvalues, and → plots them on the plane. Play with the program for large N

Hint: use SageMath

(the distribution of) Ginibre ensemble can be uniquely determined by saying that ① the joint distribution of $(\mathrm{Re}\, f_{ij}, \mathrm{Im}\, f_{ij})$ is centered Gaussian, ② specifying the covariance:

$$\mathbb{E}\, f_{ij} f_{kl} = 0$$

$$\mathbb{E}\, \overline{f_{ij}}\, \overline{f_{kl}} = 0 \qquad C = 2\delta^2$$

$$\mathbb{E}\, f_{ij}\, \overline{f_{kl}} = [i=k]\,[j=l]\, C$$

$\longrightarrow$ [MS] Section 4.3, exercise 2.

Exercise. Suppose that X is a random matrix from the Ginibre ensemble and U, V are unitary matrices.

Prove that UXV is also a random matrix from Ginibre ensemble.

<u>Hint.</u>

Calculate the covariance of the Gaussian random variables of the form

$$\langle u, Xv \rangle$$

GUE

suppose X is a random matrix from Ginibre ensemble.

we say that $Y := X + X^*$ is a

GUE (Gaussian Unitary Ensemble) random matrix.

The distribution of a GUE random matrix can be uniquely determined by saying that the entries

(g_{ij}) of Y form a complex centered Gaussian random vector with

① $g_{ij} = \overline{g_{ji}}$

② $\mathbb{E}\, g_{ij} g_{kl} = \mathbb{E}\left(f_{ij} + \overline{f_{ji}}\right)\left(f_{kl} + \overline{f_{lk}}\right) =$

$= [i=l][j=k]\ 2c$

$\boxed{2c = \frac{1}{N}}$ preferred normalization used by Mingo and Speicher.

$\boxed{2c = 1}$ another normalization used by MCS

Exercise.
Suppose Y is a GUE and
U is a unitary matrix.

Show that $U Y U^{-1}$ is a GUE.

"every orthonormal base of eigenvectors has equal probability".

Exercise. Describe the joint probability
distribution of the random variables
$(\mathrm{Re}\, g_{ij}, \mathrm{Im}\, g_{ij})_{i<j}, \ (g_{ii})_i$

Meditate about the difference between the
diagonal and the off-diagonal entries.

Exercise. Traditional computer experiment
concerning the eigenvalues.

Hint: use HISTOGRAMS

Genus expansion for GUE

$\rightarrow$ [MS] Section 1.7
Section 1.8.

$$\mathbb{E} \int_{\mathbb{R}} z^{2k} \, d\mu_Y =$$

$$= \mathbb{E}\, \mathrm{tr}\, Y^{2k} = \mathbb{E}\, \frac{1}{N} \mathrm{Tr}\, Y^{2k} = \quad \ldots$$

normalized trace

Hint:

$$\mathbb{E}\, \mathrm{tr}\, Y^{2k+1} = 0$$

for obvious reasons.

Genus expansion for GUE

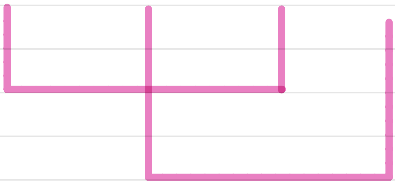
$\longrightarrow$ [MS] Section 1.7
Section 1.8.

normalized trace

γ

$$\frac{1}{N}\,\mathbb{E}\sum_{i_1,\dots,i_{2k}} g_{i_1 i_2}\, g_{i_2 i_3}\, g_{i_3 i_4}\, g_{i_4 i_5} \cdots\cdots g_{i_{2k} i_1} =$$

π



normalized trace

$$= \sum_{\pi} N^{\#\text{closed loops} \; - \; k \; - \; 1}$$

produced by covariance.

for normalization
$\mathbb{E}\,|g_{ij}|^2 = 1/N$

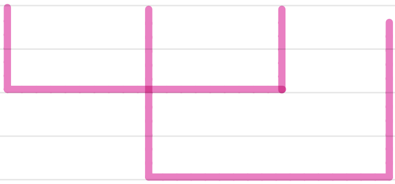
Genus expansion for GUE

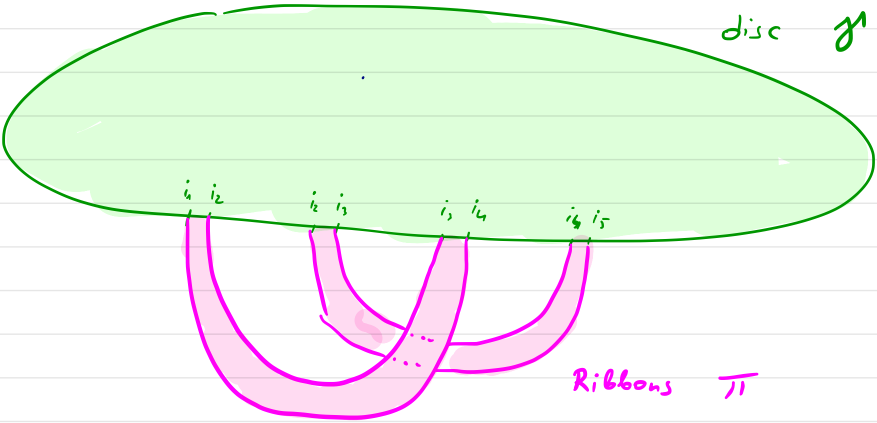
Mingo and Speicher [Section 1.8] have alternative technology.

γ

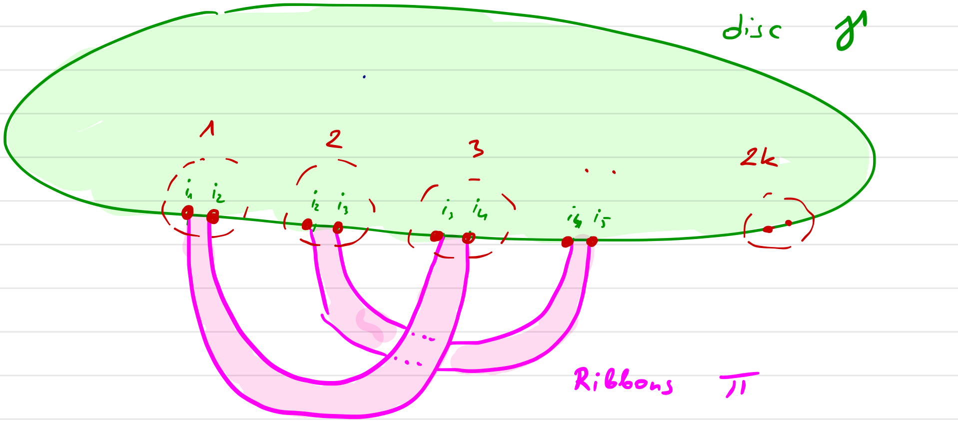
$$\mathbb{E}\sum_{i_1,\dots,i_{2k}} g_{i_1 i_2}\, g_{i_2 i_3}\, g_{i_3 i_4}\, g_{i_4 i_5} \cdots\cdots g_{i_{2k} i_1} =$$

π





oriented surface with boundary.



oriented
surface with boundary.

boundary = a number of circles S^1.
don't like surfaces with a boundary?
glue a disk to each circle S^1.

$\longrightarrow$ connected, oriented
surface without boundary.

Euler characteristic

$$\chi = 2 - 2g = V - E + F =$$

genus

$V = 4k$
$E = 4k + 2k$
$F = 1 + k + \#\text{loops}$

$$= 4k - 4k - 2k + 1 + k + \#\text{loops} =$$

$$= 1 - k + \#\text{loops}$$

$$\#\text{loops} - k - 1 = -2g$$

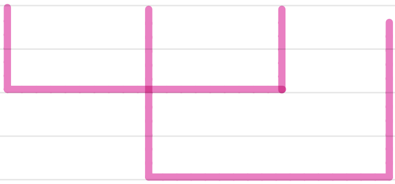
✓.

normalized trace

γ_1

$$\frac{1}{N}\;\mathbb{E}\sum_{i_1,\dots,i_{2k}} g_{i_1 i_2}\; g_{i_2 i_3}\; g_{i_3 i_4}\; g_{i_4 i_5}\;\cdots\cdots\; g_{i_{2k} i_1} =$$

π



normalized trace

$$= \sum_{\pi} N^{\#\text{closed loops} \;-\; k \;-\; 1} =$$

produced by covariance.

$$= \sum_{\pi} \frac{1}{N^{2-\chi}} = \sum_{\pi} \frac{1}{N^{2\,\mathrm{genus}(\pi)}}$$

leading contribution — genus $= 0$
sphere

requires some arguments → [MS, Section 1.8]

$$\underbrace{\lim_{n\to\infty} \mathbb{E}\, \mathrm{tr}\, Y^{2k}}_{\text{the limit exists and is finite}} = \#\text{ Non-crossing partitions on } [2k]$$

$$= C_k \quad \text{Catalan number}$$

$$= \frac{1}{k+1}\binom{2k}{k} = \quad \text{exercise.}$$

the limit exists and is finite

$$= \int_{-1}^{1} \frac{2}{\pi}\sqrt{1-x^2}\; x^{2k}\, dx =$$

this shows that the MEAN eigenvalues distribution $\mathbb{E}\,\mu_X$ converges to semicircular law μ_{sc}.
which is nice, but not exactly what we want.

$$= \int_{\mathbb{R}} x^{2k}\, d\mu_{sc}$$

$$\mathbb{E}\,\mathrm{tr}\,Y^{2k} = C_k + O\left(\frac{1}{N^2}\right)$$

<u>Exercise.</u> GOE is a ~~hermitian~~ symmetric, real valued matrix defined similarly as GUE except that ~~COMPLEX GAUSSIAN~~ is replaced by REAL GAUSSIAN.

What changes in the above calculations when GUE is replaced by GOE?

Alternative viewpoint from [MS, Section 1.8].

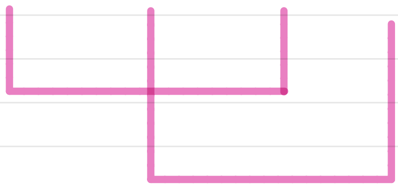
"Look on the left indices!"

γ

$$\frac{1}{N} \, \mathbb{E} \sum_{i_1,\dots,i_{2k}} \overset{1}{g_{i_1 i_2}} \, \overset{2}{g_{i_2 i_3}} \, \overset{3}{g_{i_3 i_4}} \, \overset{4}{g_{i_4 i_5}} \cdots\cdots g_{i_{2k} i_1} =$$

γ π

π



"#closed loops" =

= #cycles of $(\gamma\pi)$

$\gamma = (1, 2, \dots, 2k)$ full forward cycle

$\pi \in S_{2k}$ permutation which corresponds to the pair-partition.

$$= N^{\#\gamma\pi - k - 1}$$

$$\#\gamma\pi - k - 1 = ?$$

① length on the group S_n

$$|\sigma| = n - \#\sigma$$

the minimal number of factors to write σ as a product of transpositions.

② $|\gamma| \leq |\gamma\pi| + |\pi^{-1}|$

$$2k - \underbrace{\#\gamma}_{1} \leq 2k - \#\gamma\pi + 2k - \underbrace{\#\pi}_{k}$$

$$\boxed{\#\gamma\pi \leq k+1.}$$

when triangle inequality becomes equality?

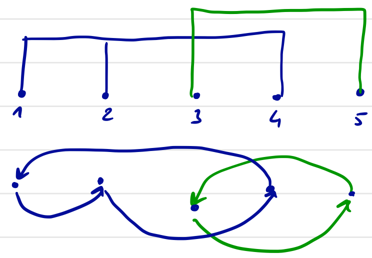
61. Robert Cori. *Un code pour les graphes planaires et ses applications.* Astérisque, No. 27. Société Mathématique de France, Paris, 1975.

Philippe Biane. Some properties of crossings and partitions. DISCRETE MATH 1997 [Section 1.3.3, Theorem 1].

Biane's result.

Given partition π of $[n] = \{1, \cdots, n\}$.

Encode π as a permutation from S_n



this is a very clever idea. We will come back to it while studying Weingarten calculus

partition π is NONCROSSING iff

$$|\gamma| = |\pi| + |\pi^{-1}\gamma|$$

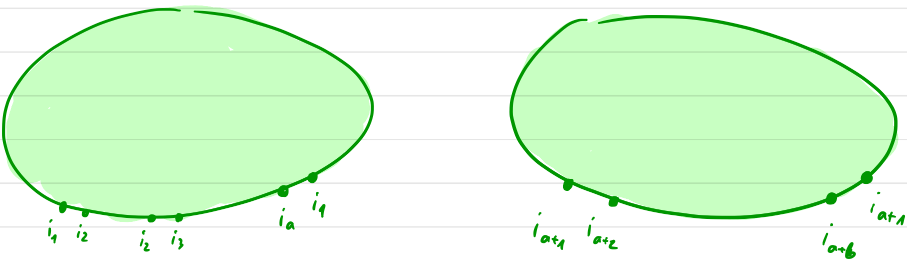
COROLLARY: we count NONCROSSING PAIR PARTITIONS

genus expansion for covariance

$$\mathbb{E}\,(\operatorname{tr} Y^a)(\operatorname{tr} Y^b) =$$

$$\frac{1}{n^2} \sum_{\pi} \sum_{i_1, \dots}$$

$$\mathbb{E}\; g_{i_1 i_2} \cdots g_{i_a i_1} \quad g_{i_{a+1} i_{a+2}}\, g_{i_{a+2} i_{a+3}} \cdots g_{i_{a+b} i_{a+1}} =$$



. . .

$$= \sum_{\pi} \frac{1}{N^{4-\chi}}$$

genus expansion for covariance

$$\mathbb{E}\left(\operatorname{tr} Y^a\right)\left(\operatorname{tr} Y^b\right) =$$

$$= \sum_{\pi} \frac{1}{N^{4-\chi}} =$$

$$= \sum_{\pi\text{-non-connected}} + \sum_{\pi\text{- connected}} \frac{1}{N^{4-\chi}}$$

$\pi = \pi_1 \cup \pi_2$

π_2: pairing of $\{a+1, a+2, \dots, a+b\}$

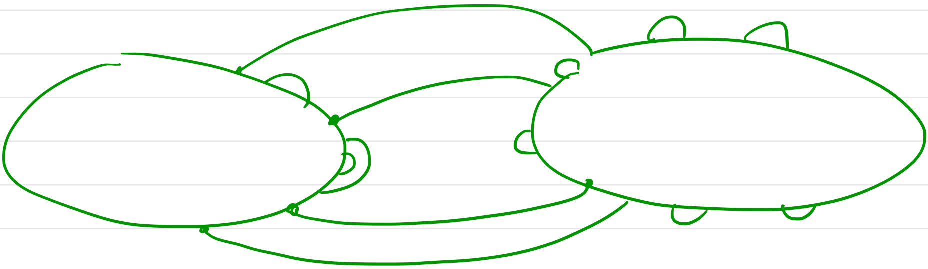
π_1: pairing of $\{1, 2, \dots, a\}$

cartesian product

$$= \underbrace{\sum_{\pi_1} \frac{1}{N^{2-\chi(\pi_1)}}}_{\mathbb{E} \operatorname{tr} Y^a} \underbrace{\sum_{\pi_2} \frac{1}{N^{2-\chi(\pi_2)}}}_{\mathbb{E} \operatorname{tr} Y^b}$$

$$\mathrm{Cov}\left(\operatorname{tr} Y^a, \operatorname{tr} Y^b\right) =$$

$$= \sum_{\substack{\pi \\ \text{CONNECTED}}} \frac{1}{N^{4-x}}$$



Minimal possible value of x = ~~4~~

= two spheres; not possible because of connectedness

= 2

= two green disks connected by a pipe = = single sphere.

$$\mathrm{Cov}\left(\operatorname{tr} Y^a, \operatorname{tr} Y^b\right) = O\left(\frac{1}{N^2}\right)$$

Corollary.

The spectral measure of a GUE random matrix converges* to the semicircular law as its size tends to infinity.

* convergence in the following senses:

→ almost surely.

this requires putting GUE random matrices into the same (product) probability space.

→ in probability.

the topology of weak convergence can be metrized. The convergence IN PROBABILITY can be formulated for metric spaces.

Hint: Chebyshev's inequality

Borel-Cantelli lemma

$$\mathbb{P}\left(|Z - \mathbb{E}Z| > t\right) \leq \frac{\operatorname{Var} Z}{t^2}$$