Random matrices and representation theory

Yi Sun

Massachusetts Institute of Technology

November 30, 2015

I. Random matrices and sample covariance matrices

II. Dynamic eigenvalues with null population covariance

III. Static eigenvalues with general population covariance

IV. Special functions from quantum groups

Random matrices — Introduction

What is a random matrix model?

- a matrix with random entries, or
- a probability measure on matrices.

Arise naturally in many contexts:

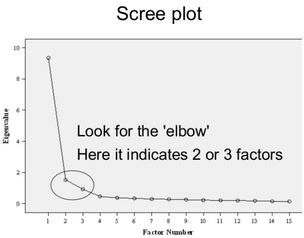
- data matrices in statistics and machine learning;
- adjacency matrices of networks in computer science;
- design matrices in compressed sensing;
- matrix models in string theory;
- scattering matrices of random media in materials science;
- statistical models for atoms with heavy nuclei.

Random matrices — Applications

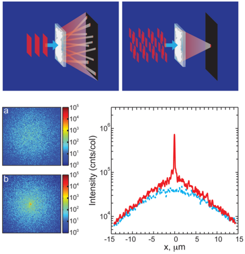
What do we study about random matrices?

- the distribution of their eigenvalues and eigenvectors, or
- integrals over the corresponding measures.

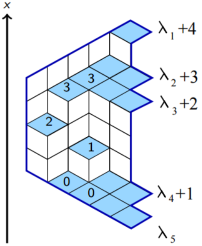
Appear in many ways: principal components, spectral gaps, coordinates for spectral clustering, transmission rates, energy levels, scaling limits of statistical mechanics models ...



Scree plot for PCA



Light in random media



Lozenge tiling

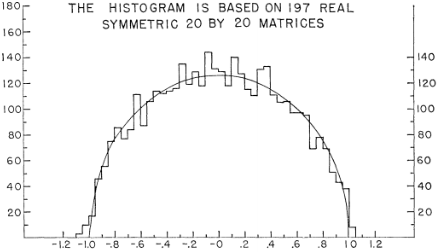
Random matrices — Examples

Basic example — Wigner ensemble:

- fix RV Y with mean 0, variance 1, and bounded moments;
- for $i \leq j$, $X_{ij} \sim Y$ and $X_{ji} = X_{ij}$;
- X symmetric $N \times N$ matrix

Histogram of eigenvalues of X for large N:

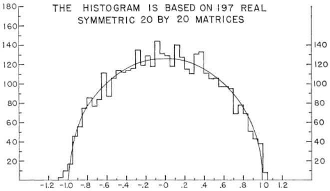
Eugene Wigner



Wigner semicircle law

Random matrices — Examples

Universality belief: limit results are **model-independent**



1. Obtain answers from exactly solvable models (this talk)
2. Prove universality to show they apply to generic models

Basic example — Gaussian Unitary Ensemble (GUE):

- X Hermitian $N \times N$ matrix, i.i.d. cplx Gaussian entries:

$$X_{ii} = N(0,2); \qquad X_{ij} = \overline{X_{ji}} = N(0,1)$$

- distribution of X invariant under unitary transformations

Unitary symmetry $\implies$ special functions $\implies$ rep. theory

Sample covariance matrices

Definition of the model:

- Let $\widehat{\Sigma}$ be the $N \times N$ population covariance
- Consider a $N \times N$ data matrix X of Gaussians
- Let X_m be the first m rows of X ($m \leq N$)

Sample covariance matrix is $\Sigma_m = X_m^* \widehat{\Sigma} X_m$

- Non-zero eigenvalues of Σ_m are:

$$0 \leq \lambda_1^{(m)} \leq \cdots \leq \lambda_m^{(m)}.$$

- Eigenvalues of Σ_m = principal values
- Eigenvectors of Σ_m = principal components

In statistics: corresponds to N variables and m samples

- Classical statistics: N fixed, $m \to \infty$.
- Modern (high-dimensional) statistics: $N, m \to \infty$ together.

Rest of the talk

First goal: Algebraic structure of sample covariance matrices

1. Dynamical construction under null population covariance
2. Density for non-null population covariance via HO functions

Second goal: Study special functions via representation theory

1. Integral formulas for Heckman-Opdam functions
2. HO functions via semiclassical limits of quantum groups

Overall goal: Link **algebraic** and **probabilistic** techniques

Laguerre eigenvalue process

Dynamic model for null population covariance ($\widehat{\Sigma} = I_{N\times N}$):

- Let $X_m(t)$ be $m \times N$ matrix of complex **Brownian motions.**
- Define $\Sigma_m(t) = X_m(t)^* X_m(t)$ with non-zero eigenvalues

$$0 \leq \lambda_1^{(m)}(t) \leq \cdots \leq \lambda_m^{(m)}(t).$$

Konig-O'Connell '01 (Laguerre eigenvalue process):

- $(\lambda_1^{(m)}(t) \leq \cdots \leq \lambda_m^{(m)}(t))$ are Markov and solve

$$d\lambda_i^{(m)} = 2\sqrt{\lambda_i^{(m)}}dB_i + 2\left(N + \sum_{j\neq i} \frac{\lambda_i^{(m)} + \lambda_j^{(m)}}{\lambda_i^{(m)} - \lambda_j^{(m)}}\right) dt$$

- Corresponds to m independent dimension $2(N-m+1)$ squared Bessel processes conditioned never to intersect

Laguerre corners process

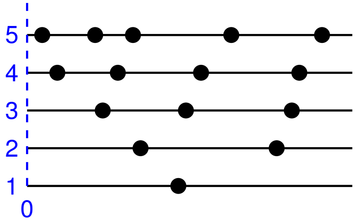
Static multilevel model for null population covariance:

- Consider $\Sigma_m = X_m^* X_m$ with non-zero eigenvalues

$$0 \le \lambda_1^{(m)} \le \cdots \le \lambda_m^{(m)}$$

for changing m.

Eigenvalues of Σ_m and Σ_{m-1} **interlace** ($\lambda_{i-1}^{(m)} \le \lambda_{i-1}^{(m-1)} \le \lambda_i^{(m)}$):



Eigenvalues form Markov chains

$$\lambda^{(1)} \to \lambda^{(2)} \to \cdots \to \lambda^{(m)}$$

and

$$\lambda^{(m)} \to \lambda^{(m-1)} \to \cdots \to \lambda^{(1)}$$

with explicit transitions.

Laguerre Warren process (I)

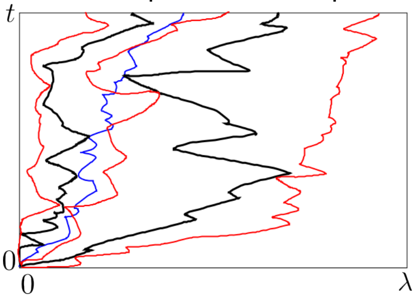
Define the Laguerre Warren process to be the solution of

$$d\lambda_i^{(m)}(t) = d\mathrm{BESQ}_i^{2(N-m+1)}(t) + dL_i^{m,+}(t) - dL_i^{m,-}(t)$$

with specified entrance law, where:

- BESQ_i^d: squared Bessel process of dim. d;
- $L_i^{m,+}$: local time at zero of $\lambda_i^{(m)} - \lambda_{i-1}^{(m-1)}$;
- $L_i^{m,-}$: local time at zero of $\lambda_i^{(m-1)} - \lambda_i^{(m)}$.

Intuition: squared Bessel processes reflecting off each other



- L1 (blue) evolves as free squared Bessel process
- L2 (black) reflects off L1
- L3 (red) reflects off L2
- ...

Laguerre Warren process (II)

Theorem (S. '15)

The Laguerre Warren process $\{\lambda_i^{(m)}\}$ is well-defined. Its projection to $(\lambda_1^{(m)}, \ldots, \lambda_m^{(m)})$ is the Laguerre eigenvalue process, and its time t marginal is the Laguerre corners process.

- Generalizes Warren '07 for GUE – replace BESQ by BM;
- S. '15: Analogous result for Jacobi ensemble – replace BESQ by univariate Jacobi process;
- Repulsion of Laguerre eigenvalue process on each level implemented by reflection on lower level.

In progress: Laguerre Warren process is limit of novel coupling of dynamics on partitions defined by Borodin-Olshanski '04.

Left edge of the Laguerre Warren process

Consider the projection to the left edge:

$$\lambda_1^{(1)} \geq \lambda_1^{(2)} \geq \cdots \geq \lambda_1^{(m)} \geq 0.$$

Evolution is Markovian and satisfies

$$d\lambda_1^{(m)}(t) = d\mathrm{BESQ}_i^{2(N-m+1)}(t) - dL_i^{m,-}(t)$$

where $L_i^{m,-}$ is local time at zero of $\lambda_1^{(m-1)} - \lambda_1^{(m)}$.

Smallest eigenvalue of sample covariance matrix $\stackrel{d}{=} \lambda_1^{(m)}(t)$

- Particle system has **local interactions**, but
- Produces hard edge of RMT!

I. Random matrices and sample covariance matrices

II. Dynamic eigenvalues with null population covariance

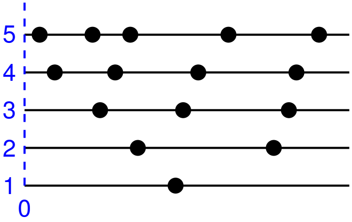
III. Static eigenvalues with general population covariance

IV. Special functions from quantum groups

General sample covariance matrices

Definition of the model:

- Take parameters $\pi_1, \ldots, \pi_N$ and $\widehat{\pi}_1, \ldots, \widehat{\pi}_N$
- Let $X_{ij} \sim N(0, (\pi_i + \widehat{\pi}_j)^{-1})$ be a **real** or **complex** Gaussian
- Define the sample covariance $\Sigma_m = X_m^* X_m$
- Eigenvalues of Σ_m interlace



Theorem (S. '15)

Eigenvalues form Markov chains

$$\lambda^{(1)} \to \lambda^{(2)} \to \cdots \to \lambda^{(m)}$$

and

$$\lambda^{(m)} \to \lambda^{(m-1)} \to \cdots \to \lambda^{(1)}$$

with explicit transitions.

Multivariate Bessel functions

We use the definitions:

$$\Delta(\lambda) = \prod_{i<j}(\lambda_i - \lambda_j) \qquad \Delta(\lambda, \mu) = \prod_{i,j}(\lambda_i - \mu_j).$$

For $\theta > 0$, define the **multivariate Bessel function** by:

$$\mathcal{B}_\theta^{N,N}(\lambda, s) = \Delta(\lambda)^{-\theta} \int_{\mu^1 \prec \cdots \prec \mu^N = \lambda} e^{\sum_{m=1}^N s_m(|\mu^m| - |\mu^{m-1}|)} \prod_{m=1}^{N-1} \frac{\Delta(\mu^m, \mu^{m+1})^{\theta-1}}{\Delta(\mu^m)^{\theta-1}\Delta(\mu^{m+1})^{\theta-1}} \prod_{m,i} d\mu_i^m.$$

- at $\theta = 1, 1/2$, recover orbital integrals for $U(N)$ and $O(N)$
- $\mathcal{B}_1^{N,N}(\lambda, s) = \frac{\det(e^{\lambda_i s_j})}{\Delta(\lambda)\Delta(s)}$ is renormalization of Schur function

Define also the conjugate by

$$\widetilde{\mathcal{B}}_\theta^{N,N}(\lambda, s) = \prod_{i=1}^N \lambda_i^{\theta-1} \mathcal{B}^{N,N}(\lambda, s).$$

Eigenvalues of general sample covariance matrices

Let $\theta = 1/2$ in the real case and $\theta = 1$ in the complex case.

Theorem (S. '15)

The joint density of $\{\lambda_i^{(m)}\}_{1\le m\le N}$ is proportional to

$$\Delta(\lambda^{(N)})^{\theta} e^{-\theta \sum_{m=1}^{N} \widehat{\pi}_m(|\lambda^{(m)}|-|\lambda^{(m-1)}|)} \prod_{m=1}^{N-1} \frac{\Delta(\lambda^{(m)},\lambda^{(m+1)})^{\theta-1}}{\Delta(\lambda^{(m)})^{\theta-1}\Delta(\lambda^{(m+1)})^{\theta-1}} \widetilde{\mathcal{B}}_{\theta}^{N,N}(\lambda^{(N)}, -\theta\pi).$$

The single-level density of $\{\lambda_i^{(N)}\}_{1\le i\le N}$ is proportional to

$$\Delta(\lambda^{(N)})^{2\theta} \mathcal{B}_{\theta}^{N,N}(\lambda^{(N)}, -\theta\widehat{\pi}) \widetilde{\mathcal{B}}_{\theta}^{N,N}(\lambda^{(N)}, -\theta\pi).$$

- Defines **multivariate Bessel measure**
- Schur measure ($\theta = 1$): Borodin-Péché, Dieker-Warren '09

Eigenvalues of general sample covariance matrices

Theorem (S. '15)

The single-level density of $\{\lambda_i^{(N)}\}_{1 \leq i \leq N}$ *is proportional to*

$$\Delta(\lambda^{(N)})^{2\theta} \mathcal{B}_\theta^{N,N}(\lambda^{(N)}, -\theta\widehat{\pi}) \widetilde{\mathcal{B}}_\theta^{N,N}(\lambda^{(N)}, -\theta\pi).$$

The $\mathcal{B}_\theta^{N,N}(\lambda, s)$ are eigenfunctions for rational Calogero-Moser:

$$H_{p_2} = \Delta(\lambda)^{\theta-1} \circ \left(\Delta + 2\theta(1-\theta) \sum_{i<j} \frac{1}{(\lambda_i - \lambda_j)^2} \right) \circ \Delta(\lambda)^{1-\theta}.$$

For symmetric p, have Hamiltonian H_p with $[H_p, H_{p_2}] = 0$ so

$$H_p \mathcal{B}_\theta^{N,N}(\lambda, s) = p(s) \mathcal{B}_\theta^{N,N}(\lambda, s).$$

Question: Asymptotics from this connection?

I. Random matrices and sample covariance matrices

II. Dynamic eigenvalues with null population covariance

III. Static eigenvalues with general population covariance

IV. Special functions from quantum groups

Eigenvalues of Jacobi ensembles

Heckman-Opdam hypergeometric functions $\mathcal{F}_\theta^{N,N}(\lambda, s)$ are eigenfunctions for **trigonometric** Calogero-Moser:

$$H_{p_2} = \Delta^{\text{trig}}(\lambda)^{\theta-1} \circ \left(\Delta + \theta(1-\theta) \sum_{i<j} \frac{1}{\sinh^2\left(\frac{\lambda_i - \lambda_j}{2}\right)} \right) \circ \Delta^{\text{trig}}(\lambda)^{1-\theta}.$$

For X a $A \times N$ Gaussian and Y_m a $m \times N$ Gaussian, define

$$Z_m = X^* X (X^* X + Y_m^* Y_m)^{-1}.$$

Theorem (S. '15)

For $\pi = (A - N + 1, \ldots, A)$ and $\widehat{\pi} = (0, 1, \ldots, N-1)$, the non 0-1 eigenvalues $\mu_i^{(m)}$ of Z_m have single level density

$$\Delta^{trig}(\lambda^{(N)})^{2\theta} \mathcal{F}_\theta^{N,N}(\lambda^{(N)}, -\theta\pi) \widetilde{\mathcal{F}}_\theta^{N,N}(\lambda^{(N)}, -\theta\widehat{\pi})$$

and multilevel density determined by branching.

Eigenvalues of Jacobi ensembles

Theorem (S. '15)

For $\pi = (A - N + 1, \dots, A)$ *and* $\widehat{\pi} = (0, 1, \dots, N-1)$*, the non 0-1 eigenvalues* $\mu_i^{(m)}$ *of* Z_m *have single level density*

$$\Delta^{trig}(\lambda^{(N)})^{2\theta} \mathcal{F}_\theta^{N,N}(\lambda^{(N)}, -\theta\pi) \widetilde{\mathcal{F}}_\theta^{N,N}(\lambda^{(N)}, -\theta\widehat{\pi})$$

and multilevel density determined by branching.

- Borodin-Gorin '13: Showed this multilevel HO process converges to a 2-D Gaussian free field
- Result gives a link to random matrices
- Techniques of Borodin-Gorin '13 based on degeneration of formulas from Macdonald processes of Borodin-Corwin '11

Heckman-Opdam hypergeometric functions

Borodin-Gorin '13: Integral formula for HO function:

$$\mathcal{F}_\theta^{N,N}(\lambda, s) = C_N \Delta^{\text{trig}}(\lambda)^{-\theta} \int_{\mu^1 \prec \cdots \prec \mu^N = \lambda} e^{\sum_{m=1}^N s_m(|\mu^m| - |\mu^{m-1}|)}$$

$$\prod_{m=1}^{N-1} \frac{\Delta(e^{\mu^m}, e^{\mu^{m+1}})^{\theta-1}}{\Delta(e^{\mu^m})^{\theta-1} \Delta(e^{\mu^{m+1}})^{\theta-1}} e^{-(\theta-1)|\mu^m|} \prod_i d\mu_i^m,$$

where $\Delta^{\text{trig}}(\lambda) = \prod_{i<j} \left(e^{\frac{\lambda_i - \lambda_j}{2}} - e^{\frac{\lambda_j - \lambda_i}{2}} \right)$.

Proof: Take quasiclassical limit of quantum story.

- In this setting, **quantum** $=$ **discrete**
- Rest of the talk: Explain what this means algebraically.

Macdonald polynomials

Macdonald '88: Introduced symmetric polynomials

$$P_\lambda(x_1, \ldots, x_N; q, t)$$

indexed by partitions λ. Example:

$$P_{(3)}(x, y; t, q) = x^3 + y^3 + \frac{1 - t + q - qt + q^2 - tq^2}{1 - tq^2}(x^2y + y^2x).$$

Generalizes many special functions: Schur, Heckman-Opdam

For partition λ, $P_\lambda(x; q, t)$ is joint eigenfunction of:

- q-difference operators $D_n^r(q, t)$ in x
- with eigenvalue $e_r(q^\lambda t^\rho)$ depending on λ.

Quasiclassical limit of formulas

In quasiclassical limit, as $\varepsilon \to 0$:

$$q = e^{-\varepsilon} \qquad t = e^{-\theta\varepsilon} \qquad x = e^{\varepsilon s} \qquad \lambda = \varepsilon^{-1}\Lambda.$$

HO function is quasiclassical limit of Macdonald polynomial:

$$\lim_{\varepsilon\to 0} \frac{P_{\lfloor \varepsilon^{-1}(\lambda_1,\ldots,\lambda_N)\rfloor}(e^{\varepsilon s_1},\ldots,e^{\varepsilon s_N}; e^{-\varepsilon}, e^{-\theta\varepsilon})}{P_{\lfloor \varepsilon^{-1}(\lambda_1,\ldots,\lambda_N)\rfloor}(1,\ldots,e^{-(N-1)\theta\varepsilon}; e^{-\varepsilon}, e^{-\theta\varepsilon})} = e^{\frac{n-1}{2}\theta|\lambda|}\mathcal{F}_\theta^{N,N}(\lambda, s)$$

Borodin-Gorin formula is limit of **Macdonald branching rule**:

$$P_\lambda(x_1,\ldots,x_n; q,t) = \sum_{\mu\prec\lambda} \psi_{\lambda/\mu}(q,t) P_\mu(x_1,\ldots,x_{n-1}; q,t) x_n^{|\lambda|-|\mu|}$$

for branching coefficient $\psi_{\lambda/\mu}(q,t)$ given by

$$\prod_{1\le i\le j\le \ell(\mu)} \frac{(q^{\mu_i-\mu_j}t^{j-i+1};q)(q^{\lambda_i-\lambda_{j+1}}t^{j-i+1};q)(q^{\lambda_i-\mu_j+1}t^{j-i};q)(q^{\mu_i-\lambda_{j+1}+1}t^{j-i};q)}{(q^{\mu_i-\mu_j+1}t^{j-i};q)(q^{\lambda_i-\lambda_{j+1}+1}t^{j-i};q)(q^{\lambda_i-\mu_j}t^{j-i+1};q)(q^{\mu_i-\lambda_{j+1}}t^{j-i+1};q)}.$$

Etingof-Kirillov Jr. approach

For each λ, can construct linear maps (**intertwiners**)

$$\Phi_\lambda^N : L_{\lambda+(\theta-1)\rho} \to L_{\lambda+(\theta-1)\rho} \otimes W_{\theta-1}$$

between representations of the **quantum group** $U_q(\mathfrak{gl}_N)$.

Theorem (Etingof-Kirillov Jr. '94)

For integer θ, the Macdonald polynomial $P_\lambda(x; q^2, q^{2\theta})$ is

$$P_\lambda(x; q^2, q^{2\theta}) = \frac{Tr(\Phi_\lambda^N x^h)}{Tr(\Phi_0^N x^h)}.$$

Note: Interpret traces of Φ_λ^N as scalars via $W_{\theta-1}[0] \simeq \mathbb{C} \cdot w_{\theta-1}$.

Quasiclassical limit of algebras

Theorem (S. '15)

If θ is a positive integer, the HO function $\mathcal{F}_\theta^{N,N}(\lambda, s)$ is given by:

$$\mathcal{F}_\theta^{N,N}(\lambda, s) = C_N \frac{1}{\Delta^{trig}(\lambda)^\theta \prod_{a=1}^{\theta-1} \prod_{i<j}(s_i - s_j - a)} \int_{X \in \mathcal{O}_\Lambda} F_{\theta-1}(X) \prod_{m=1}^{N} \left(\frac{\det(X_m)}{\det(X_{m-1})} \right)^{s_m} d\mu_\Lambda.$$

- X_m is the principal $m \times m$ submatrix of X;
- $\mathcal{O}_\Lambda$ is the **dressing orbit** of $U(N)$ on $B(N)$ containing e^λ;
- $F_{\theta-1} : \mathcal{O}_\Lambda \to W_{\theta-1}$ is $U(N)$-invariant.

Proof by general result on quasiclassical limits:

$$\lim_{\varepsilon \to 0} \mathrm{Tr}|_{L_{\varepsilon^{-1}\lambda}}(-) \mapsto \int_{\mathcal{O}_\Lambda} - d\mu_\Lambda, \qquad \Phi_\lambda^N \mapsto F_{\theta-1}$$

Summary

This talk:

1. Random matrices and sample covariance matrices
2. Dynamic eigenvalues with null population covariance
3. Static eigenvalues with general population covariance
4. Macdonald and HO functions via quantum groups

References:

- Y. S. A Laguerre analogue of the Warren process, in preparation, 2015.
- Y. S. Matrix models for multilevel Heckman-Opdam and multivariate Bessel measures, preprint, 2015.
- Y. S. A new integral formula for the Heckman-Opdam hypergeometric functions, Advances in Mathematics, to appear.
- Y. S. A representation-theoretic proof of the branching rule for Macdonald polynomials, Mathematical Research Letters, to appear.