Random Unitary Matrices and Friends

Elizabeth Meckes

Case Western Reserve University

LDHD Summer School
SAMSI
August, 2013

What is a random unitary matrix?

What is a random unitary matrix?

- A unitary matrix is an $n \times n$ matrix U with entries in $\mathbb{C}$, such that

$$UU^* = I,$$

where U^* is the conjugate transpose of U.

What is a random unitary matrix?

- A unitary matrix is an $n \times n$ matrix U with entries in $\mathbb{C}$, such that

$$UU^* = I,$$

where U^* is the conjugate transpose of U.

That is, a unitary matrix is an $n \times n$ matrix over $\mathbb{C}$ whose columns (or rows) are orthonormal in $\mathbb{C}^n$.

What is a random unitary matrix?

- A unitary matrix is an $n \times n$ matrix U with entries in $\mathbb{C}$, such that
$$UU^* = I,$$
where U^* is the conjugate transpose of U.

  That is, a unitary matrix is an $n \times n$ matrix over $\mathbb{C}$ whose columns (or rows) are orthonormal in $\mathbb{C}^n$.

- The set of all $n \times n$ unitary matrices is denoted $\mathbb{U}(n)$; this set is a group and a manifold.

What is a random unitary matrix?

- Metric Structure:
  - $\mathbb{U}(n)$ sits inside $\mathbb{C}^{n^2}$ and inherits a geodesic metric $d_g(\cdot, \cdot)$ from the Euclidean metric on $\mathbb{C}^{n^2}$.

What is a random unitary matrix?

- Metric Structure:
  - $\mathbb{U}(n)$ sits inside $\mathbb{C}^{n^2}$ and inherits a geodesic metric $d_g(\cdot, \cdot)$ from the Euclidean metric on $\mathbb{C}^{n^2}$.
  - $\mathbb{U}(n)$ also has its own Euclidean (Hilbert-Schmidt) metric from the inner product $\langle U, V\rangle = \mathrm{Tr}(UV^*)$.

What is a random unitary matrix?

- Metric Structure:
  - $\mathbb{U}(n)$ sits inside $\mathbb{C}^{n^2}$ and inherits a geodesic metric $d_g(\cdot,\cdot)$ from the Euclidean metric on $\mathbb{C}^{n^2}$.
  - $\mathbb{U}(n)$ also has its own Euclidean (Hilbert-Schmidt) metric from the inner product $\langle U, V \rangle = \mathrm{Tr}(UV^*)$.
  - The two metrics are equivalent:

$$d_{HS}(U, V) \leq d_g(U, V) \leq \frac{\pi}{2} d_{HS}(U, V).$$

What is a random unitary matrix?

- Metric Structure:
  - $\mathbb{U}(n)$ sits inside $\mathbb{C}^{n^2}$ and inherits a geodesic metric $d_g(\cdot, \cdot)$ from the Euclidean metric on $\mathbb{C}^{n^2}$.
  - $\mathbb{U}(n)$ also has its own Euclidean (Hilbert-Schmidt) metric from the inner product $\langle U, V \rangle = \mathrm{Tr}(UV^*)$.
  - The two metrics are equivalent:

$$d_{HS}(U, V) \leq d_g(U, V) \leq \frac{\pi}{2} d_{HS}(U, V).$$

- Randomness:
  There is a unique translation-invariant probability measure called Haar measure on $\mathbb{U}(n)$: if U is a Haar-distributed random unitary matrix, so are AU and UA, for A a fixed unitary matrix.

A couple ways to build a random unitary matrix

A couple ways to build a random unitary matrix

1. ▸ Pick the first column U_1 uniformly from $\mathbb{S}^1_{\mathbb{C}} \subseteq \mathbb{C}^n$.

A couple ways to build a random unitary matrix

1.
   - Pick the first column U_1 uniformly from $\mathbb{S}^1_{\mathbb{C}} \subseteq \mathbb{C}^n$.
   - Pick the second column U_2 uniformly from $U_1^{\perp} \subseteq \mathbb{S}^1_{\mathbb{C}}$.

A couple ways to build a random unitary matrix

1.
   - Pick the first column U_1 uniformly from $\mathbb{S}^1_{\mathbb{C}} \subseteq \mathbb{C}^n$.
   - Pick the second column U_2 uniformly from $U_1^\perp \subseteq \mathbb{S}^1_{\mathbb{C}}$.

     $\vdots$
   - Pick the last column U_n uniformly from $(span\{U_1, \ldots, U_{n-1}\})^\perp \subseteq \mathbb{S}^1_{\mathbb{C}}$.

A couple ways to build a random unitary matrix

1.
   - Pick the first column U_1 uniformly from $\mathbb{S}^1_{\mathbb{C}} \subseteq \mathbb{C}^n$.
   - Pick the second column U_2 uniformly from $U_1^\perp \subseteq \mathbb{S}^1_{\mathbb{C}}$.

     $\vdots$
   - Pick the last column U_n uniformly from $(span\{U_1, \ldots, U_{n-1}\})^\perp \subseteq \mathbb{S}^1_{\mathbb{C}}$.

2.
   - Fill an $n \times n$ array with i.i.d. standard complex Gaussian random variables.

A couple ways to build a random unitary matrix

1.
   - Pick the first column U_1 uniformly from $\mathbb{S}^1_{\mathbb{C}} \subseteq \mathbb{C}^n$.
   - Pick the second column U_2 uniformly from $U_1^{\perp} \subseteq \mathbb{S}^1_{\mathbb{C}}$.

     ⋮
   - Pick the last column U_n uniformly from $(span\{U_1, \dots, U_{n-1}\})^{\perp} \subseteq \mathbb{S}^1_{\mathbb{C}}$.

2.
   - Fill an $n \times n$ array with i.i.d. standard complex Gaussian random variables.
   - Stick the result into the *QR* algorithm; the resulting Q is Haar-distributed on $\mathbb{U}(n)$.

Meet $\mathbb{U}(n)$'s kid sister: The orthogonal group

Meet $\mathbb{U}(n)$'s kid sister: The orthogonal group

- An orthogonal matrix is an $n \times n$ matrix U with entries in $\mathbb{R}$, such that
$$UU^T = I,$$
where U^T is the transpose of U.

Meet $\mathbb{U}(n)$'s kid sister: The orthogonal group

- An orthogonal matrix is an $n \times n$ matrix U with entries in $\mathbb{R}$, such that
$$UU^T = I,$$
where U^T is the transpose of U. That is, a unitary matrix is an $n \times n$ matrix over $\mathbb{R}$ whose columns (or rows) are orthonormal in $\mathbb{R}^n$.

Meet $\mathbb{U}(n)$'s kid sister: The orthogonal group

- An orthogonal matrix is an $n \times n$ matrix U with entries in $\mathbb{R}$, such that
$$UU^T = I,$$
where U^T is the transpose of U. That is, a unitary matrix is an $n \times n$ matrix over $\mathbb{R}$ whose columns (or rows) are orthonormal in $\mathbb{R}^n$.

- The set of all $n \times n$ unitary matrices is denoted $\mathbb{O}(n)$; this set is a subgroup and a submanifold of $\mathbb{U}(n)$.

Meet $\mathbb{U}(n)$'s kid sister: The orthogonal group

- An orthogonal matrix is an $n \times n$ matrix U with entries in $\mathbb{R}$, such that
$$UU^T = I,$$
where U^T is the transpose of U. That is, a unitary matrix is an $n \times n$ matrix over $\mathbb{R}$ whose columns (or rows) are orthonormal in $\mathbb{R}^n$.

- The set of all $n \times n$ unitary matrices is denoted $\mathbb{O}(n)$; this set is a subgroup and a submanifold of $\mathbb{U}(n)$.

- $\mathbb{O}(n)$ has two connected components: $\mathbb{SO}(n)$ ($\det(U) = 1$) and $\mathbb{SO}^-(n)$ ($\det(U) = -1$).

- There is a unique translation-invariant (Haar) probability measure on each of $\mathbb{O}(n)$, $\mathbb{SO}(n)$ and $\mathbb{SO}^-(n)$.

The symplectic group:
the weird uncle no one talks about

The symplectic group: the weird uncle no one talks about

- A symplectic matrix is an $2n \times 2n$ matrix with entries in $\mathbb{C}$, such that

$$UJU^* = J,$$

where U^* is the conjugate transpose of U and

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}.$$

The symplectic group: the weird uncle no one talks about

- A symplectic matrix is an $2n \times 2n$ matrix with entries in $\mathbb{C}$, such that

$$UJU^* = J,$$

where U^* is the conjugate transpose of U and

$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}.$$

(It is really the quaternionic unitary group.)

The symplectic group: the weird uncle no one talks about

- A symplectic matrix is an $2n \times 2n$ matrix with entries in $\mathbb{C}$, such that
$$UJU^* = J,$$
where U^* is the conjugate transpose of U and
$$J = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix}.$$
(It is really the quaternionic unitary group.)

- The group of $2n \times 2n$ symplectic matrices is denoted $\mathbb{Sp}(2n)$.

Concentration of measure

Theorem (G/M;B/E;L;M/M)

Let G be one of $\mathbb{SO}(n)$, $\mathbb{SO}^{-}(n)$, $\mathbb{SU}(n)$, $\mathbb{U}(n)$, $\mathbb{S}\mathrm{p}(2n)$, and let $F : G \to \mathbb{R}$ be L-Lipschitz (w.r.t. the geodesic metric or the HS-metric). Let U be distributed according to Haar measure on G. Then there are universal constants C, c such that

$$\mathbb{P}\left[\left|F(U) - \mathbb{E}F(U)\right| > Lt\right] \leq Ce^{-cnt^2},$$

for every $t > 0$.

The entries of a random orthogonal matrix

Note: permuting the rows or columns of a random orthogonal matrix U corresponds to left- or right-multiplication by a permutation matrix (which is itself orthogonal).

The entries of a random orthogonal matrix

Note: permuting the rows or columns of a random orthogonal matrix U corresponds to left- or right-multiplication by a permutation matrix (which is itself orthogonal).

$\Longrightarrow$ The entries $\{u_{ij}\}$ of U all have the same distribution.

The entries of a random orthogonal matrix

Note: permuting the rows or columns of a random orthogonal matrix U corresponds to left- or right-multiplication by a permutation matrix (which is itself orthogonal).

$\Longrightarrow$ The entries $\{u_{ij}\}$ of U all have the same distribution.

Classical fact: A coordinate of a random point on the sphere in $\mathbb{R}^n$ is approximately Gaussian, for large n.

The entries of a random orthogonal matrix

Note: permuting the rows or columns of a random orthogonal matrix U corresponds to left- or right-multiplication by a permutation matrix (which is itself orthogonal).

$\Longrightarrow$ The entries $\{u_{ij}\}$ of U all have the same distribution.

Classical fact: A coordinate of a random point on the sphere in $\mathbb{R}^n$ is approximately Gaussian, for large n.

$\Longrightarrow$ The entries $\{u_{ij}\}$ of U are

individually approximately Gaussian

if U is large.

The entries of a random orthogonal matrix

A more modern fact (Diaconis–Freedman): If X is a randomly distributed point on the sphere of radius $\sqrt{n}$ in $\mathbb{R}^n$, and Z is a standard Gaussian random vector in $\mathbb{R}^n$, then

$$d_{TV}\Big((X_1, \ldots, X_k), (Z_1, \ldots, Z_k)\Big) \leq \frac{2(k+3)}{n-k-3}.$$

The entries of a random orthogonal matrix

A more modern fact (Diaconis–Freedman): If X is a randomly distributed point on the sphere of radius $\sqrt{n}$ in $\mathbb{R}^n$, and Z is a standard Gaussian random vector in $\mathbb{R}^n$, then

$$d_{TV}\Big((X_1, \ldots, X_k), (Z_1, \ldots, Z_k)\Big) \leq \frac{2(k+3)}{n-k-3}.$$

$\Longrightarrow$ Any k entries within one row (or column) of $U \in \mathbb{U}(n)$ are approximately independent Gaussians, if $k = o(n)$.

The entries of a random orthogonal matrix

A more modern fact (Diaconis–Freedman): If X is a randomly distributed point on the sphere of radius $\sqrt{n}$ in $\mathbb{R}^n$, and Z is a standard Gaussian random vector in $\mathbb{R}^n$, then

$$d_{TV}\Big((X_1,\ldots,X_k),(Z_1,\ldots,Z_k)\Big) \le \frac{2(k+3)}{n-k-3}.$$

$\Longrightarrow$ Any k entries within one row (or column) of $U \in \mathbb{U}(n)$ are approximately independent Gaussians, if $k = o(n)$.

Diaconis'question: How many entries of U can be simultaneously approximated by independent Gaussians?

Jiang's answer(s)

Jiang's answer(s)

It depends on what you mean by approximated.

Jiang's answer(s)

It depends on what you mean by approximated.

Theorem (Jiang)

Let $\{U_n\}$ be a sequence of random orthogonal matrices with $U_n \in \mathbb{O}(n)$ for each n, and suppose that $p_n, q_n = o(\sqrt{n})$.

Let $\mathcal{L}(\sqrt{n}U(p_n, q_n))$ denote the joint distribution of the $p_n q_n$ entries of the top-left $p_n \times q_n$ block of $\sqrt{n}U_n$, and let $Z(p_n, q_n)$ denote a collection of $p_n q_n$ i.i.d. standard normal random variables. Then

$$\lim_{n\to\infty} d_{TV}(\mathcal{L}(\sqrt{n}U(p_n, q_n)), Z(p_n, q_n)) = 0.$$

Jiang's answer(s)

It depends on what you mean by approximated.

Theorem (Jiang)

Let $\{U_n\}$ be a sequence of random orthogonal matrices with $U_n \in \mathbb{O}(n)$ for each n, and suppose that $p_n, q_n = o(\sqrt{n})$.

Let $\mathcal{L}(\sqrt{n}U(p_n, q_n))$ denote the joint distribution of the $p_n q_n$ entries of the top-left $p_n \times q_n$ block of $\sqrt{n}U_n$, and let $Z(p_n, q_n)$ denote a collection of $p_n q_n$ i.i.d. standard normal random variables. Then

$$\lim_{n\to\infty} d_{TV}(\mathcal{L}(\sqrt{n}U(p_n, q_n)), Z(p_n, q_n)) = 0.$$

That is, a $p_n \times q_n$ principle submatrix can be approximated in total variation by a Gaussian random matrix, as long as $p_n, q_n \ll \sqrt{n}$.

Jiang's answer(s)

Theorem (Jiang)

For each n, let $Y_n = \left[y_{ij}\right]_{i,j=1}^{n}$ *be an* $n \times n$ *matrix of independent standard Gaussian random variables and let* $\Gamma_n = \left[\gamma_{ij}\right]_{i,j=1}^{n}$ *be the matrix obtained from* Y_n *by performing the Gram-Schmidt process; i.e.,* Γ_n *is a random orthogonal matrix. Let*

$$\epsilon_n(m) = \max_{1 \le i \le n, 1 \le j \le m} \left| \sqrt{n}\gamma_{ij} - y_{ij} \right|.$$

Then

$$\epsilon_n(m_n) \xrightarrow[n \to \infty]{\mathbb{P}} 0$$

if and only if $m_n = o\left(\frac{n}{\log(n)}\right)$.

Jiang's answer(s)

Theorem (Jiang)

For each n, let $Y_n = [y_{ij}]_{i,j=1}^{n}$ be an $n \times n$ matrix of independent standard Gaussian random variables and let $\Gamma_n = [\gamma_{ij}]_{i,j=1}^{n}$ be the matrix obtained from Y_n by performing the Gram-Schmidt process; i.e., Γ_n is a random orthogonal matrix. Let

$$\epsilon_n(m) = \max_{1 \le i \le n, 1 \le j \le m} \left| \sqrt{n}\gamma_{ij} - y_{ij} \right|.$$

Then

$$\epsilon_n(m_n) \xrightarrow[n\to\infty]{\mathbb{P}} 0$$

if and only if $m_n = o\left(\frac{n}{\log(n)}\right)$.

That is, in an "in probability" sense, $\frac{n^2}{\log(n)}$ entries of U can be simultaneously approximated by independent Gaussians.

A more geometric viewpoint

A more geometric viewpoint

Choosing a principle submatrix of an $n \times n$ orthogonal matrix U corresponds to a particular type of orthogonal projection from a large matrix space to a smaller one.

A more geometric viewpoint

Choosing a principle submatrix of an $n \times n$ orthogonal matrix U corresponds to a particular type of orthogonal projection from a large matrix space to a smaller one.

(Note that the result is no longer orthogonal.)

A more geometric viewpoint

Choosing a principle submatrix of an $n \times n$ orthogonal matrix U corresponds to a particular type of orthogonal projection from a large matrix space to a smaller one.

(Note that the result is no longer orthogonal.)

In general, a rank k orthogonal projection of $\mathbb{O}(n)$ looks like

$$U \mapsto \big(\mathrm{Tr}(A_1 U), \ldots, \mathrm{Tr}(A_k U) \big),$$

where $A_1, \ldots, A_k$ are orthonormal matrices in $\mathbb{O}(n)$; i.e.,

$$\mathrm{Tr}(A_i A_j^T) = \delta_{ij}.$$

A more geometric viewpoint

Theorem (Chatterjee–M.)

Let $A_1, \ldots, A_k$ be orthonormal (w.r.t. the Hilbert-Schmidt inner product) in $\mathbb{O}(n)$, and let $U \in \mathbb{O}(n)$ be a random orthogonal matrix. Consider the random vector

$$X := (\mathrm{Tr}(A_1 U), \ldots, \mathrm{Tr}(A_k U)),$$

and let $Z := (Z_1, \ldots, Z_k)$ be a standard Gaussian random vector in $\mathbb{R}^k$. Then for all $n \geq 2$,

$$d_W(X, Z) \leq \frac{\sqrt{2}k}{n-1}.$$

Here, $d_W(\cdot, \cdot)$ denotes the L_1-Wasserstein distance.

Eigenvalues – The empirical spectral measure

Eigenvalues – The empirical spectral measure

Let U be a Haar-distributed matrix in $\mathbb{U}(N)$.

Then U has (random) eigenvalues $\{e^{i\theta_j}\}_{j=1}^{N}$.

Eigenvalues – The empirical spectral measure

Let U be a Haar-distributed matrix in $\mathbb{U}(N)$.

Then U has (random) eigenvalues $\{e^{i\theta_j}\}_{j=1}^N$.

Eigenvalues – The empirical spectral measure

Let U be a Haar-distributed matrix in $\mathbb{U}(N)$.

Then U has (random) eigenvalues $\{e^{i\theta_j}\}_{j=1}^{N}$.

Note: The distribution of the set of eigenvalues is rotation-invariant.

Eigenvalues – The empirical spectral measure

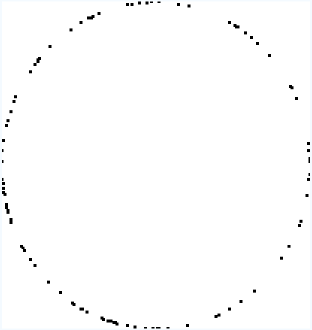
Let U be a Haar-distributed matrix in $\mathbb{U}(N)$.

Then U has (random) eigenvalues $\{e^{i\theta_j}\}_{j=1}^{N}$.

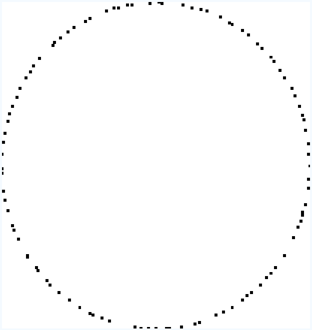
Note: The distribution of the set of eigenvalues is
rotation-invariant.

To understand the behavior of the ensemble of random eigenvalues, we consider the empirical spectral measure of U:

$$\mu_N := \frac{1}{N} \sum_{j=1}^{N} \delta_{e^{i\theta_j}}.$$



100 i.i.d. uniform random points



E. Rains

The eigenvalues of a 100×100 random unitary matrix

Diaconis/Shahshahani

Theorem (D–S)

Let $U_n \in \mathbb{U}(n)$ be a random unitary matrix, and let μ_{U_n} denote the empirical spectral measure of U_n. Let ν denote the uniform probability measure on $\mathbb{S}^1$. Then

$$\mu_{U_n} \xrightarrow{n\to\infty} \nu,$$

weak- in probability.*

Diaconis/Shahshahani

Theorem (D–S)

Let $U_n \in \mathbb{U}(n)$ be a random unitary matrix, and let μ_{U_n} denote the empirical spectral measure of U_n. Let ν denote the uniform probability measure on $\mathbb{S}^1$. Then

$$\mu_{U_n} \xrightarrow{n\to\infty} \nu,$$

weak- in probability.*

- The theorem follows from explicit formulae for the mixed moments of the random vector $\left(\mathrm{Tr}(U_n), \ldots, \mathrm{Tr}(U_n^k)\right)$ for fixed k, which have been useful in many other contexts.

Diaconis/Shahshahani

Theorem (D–S)

Let $U_n \in \mathbb{U}(n)$ be a random unitary matrix, and let μ_{U_n} denote the empirical spectral measure of U_n. Let ν denote the uniform probability measure on $\mathbb{S}^1$. Then

$$\mu_{U_n} \xrightarrow{n\to\infty} \nu,$$

weak- in probability.*

- The theorem follows from explicit formulae for the mixed moments of the random vector $\big(\mathrm{Tr}(U_n), \ldots, \mathrm{Tr}(U_n^k)\big)$ for fixed k, which have been useful in many other contexts.
- They showed in particular that $\big(\mathrm{Tr}(U_n), \ldots, \mathrm{Tr}(U_n^k)\big)$ is asymptotically distributed as a standard complex Gaussian random vector.

The number of eigenvalues in an arc

Theorem (Wieand)

Let $I_j := (e^{i\alpha_j}, e^{i\beta_j})$ be intervals on $\mathbb{S}^1$ and for $U_n \in \mathbb{U}(n)$ a random unitary matrix, let

$$Y_{n,k} := \frac{\mu_{U_n}(I_k) - \mathbb{E}\mu_{U_n}(I_k)}{\frac{1}{\pi}\sqrt{\log(n)}}.$$

Then as n tends to infinity, the random vector $(Y_{n,1}, \ldots, Y_{n,k})$ converges in distribution to a jointly Gaussian random vector $(Z_1, \ldots, Z_k)$ with covariance

$$\mathrm{Cov}(Z_j, Z_k) = \begin{cases} 0, & \alpha_j, \alpha_k, \beta_j, \beta_k \text{ all distict}; \\ \frac{1}{2} & \alpha_j = \alpha_k \text{ or } \beta_j = \beta_k \text{ (but not both)}; \\ -\frac{1}{2} & \alpha_j = \beta_k \text{ or } \beta_j = \alpha_k \text{ (but not both)}; \\ 1 & \alpha_j = \alpha_k \text{ and } \beta_j = \beta_k; \\ -1 & \alpha_j = \beta_k \text{ and } \beta_j = \alpha_k. \end{cases}$$

About that weird covariance structure...

About that weird covariance structure...

Another Gaussian process that has it:

About that weird covariance structure...

Another Gaussian process that has it: Again suppose that $I_j := (e^{i\alpha_j}, e^{i\beta_j})$ are intervals on $\mathbb{S}^1$, and suppose that $\{G_\theta\}_{\theta\in[0,2\pi)}$ are i.i.d. standard Gaussians. Define

$$X_{n,k} = G_{\beta_k} - G_{\alpha_k};$$

then

$$\mathrm{Cov}(X_j, X_k) = \begin{cases} 0, & \alpha_j, \alpha_k, \beta_j, \beta_k \text{ all distict}; \\ \frac{1}{2} & \alpha_j = \alpha_k \text{ or } \beta_j = \beta_k \text{ (but not both)}; \\ -\frac{1}{2} & \alpha_j = \beta_k \text{ or } \beta_j = \alpha_k \text{ (but not both)}; \\ 1 & \alpha_j = \alpha_k \text{ and } \beta_j = \beta_k; \\ -1 & \alpha_j = \beta_k \text{ and } \beta_j = \alpha_k. \end{cases}$$

Where's the white noise in U?

Where's the white noise in U?

Theorem (Hughes–Keating–O'Connel)

Let $Z(\theta)$ be the characteristic polynomial of U and fix $\theta_1 \ldots, \theta_k$. Then

$$\frac{1}{\sqrt{\frac{1}{2}\log(n)}}\big(\log(Z(\theta_1)), \ldots, \log(Z(\theta_k))\big)$$

converges in distribution to a standard Gaussian random vector in $\mathbb{C}^k$, as $n \to \infty$.

Where's the white noise in U?

Theorem (Hughes–Keating–O'Connel)

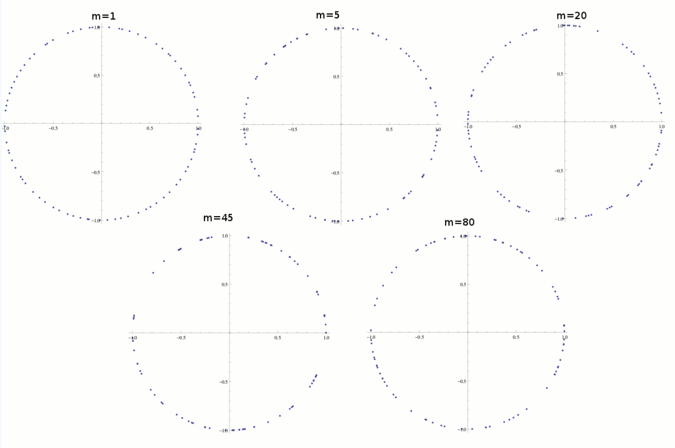
Let $Z(\theta)$ be the characteristic polynomial of U and fix $\theta_1 \ldots, \theta_k$. Then

$$\frac{1}{\sqrt{\frac{1}{2}\log(n)}}\big(\log(Z(\theta_1)), \ldots, \log(Z(\theta_k))\big)$$

converges in distribution to a standard Gaussian random vector in $\mathbb{C}^k$, as $n \to \infty$.

HKO in particular showed that Wieand's result follows from theirs by the argument principle.

Powers of U



The eigenvalues of U^m for $m = 1, 5, 20, 45, 80$, for U a realization of a random 80×80 unitary matrix.

Rains’ Theorems

Rains' Theorems

Theorem (Rains 1997)

Let $U \in \mathbb{U}(n)$ be a random unitary matrix, and let $m \geq n$. Then the eigenvalues of U^m are distributed exactly *as n i.i.d. uniform points on $\mathbb{S}^1$.*

Rains' Theorems

Theorem (Rains 1997)

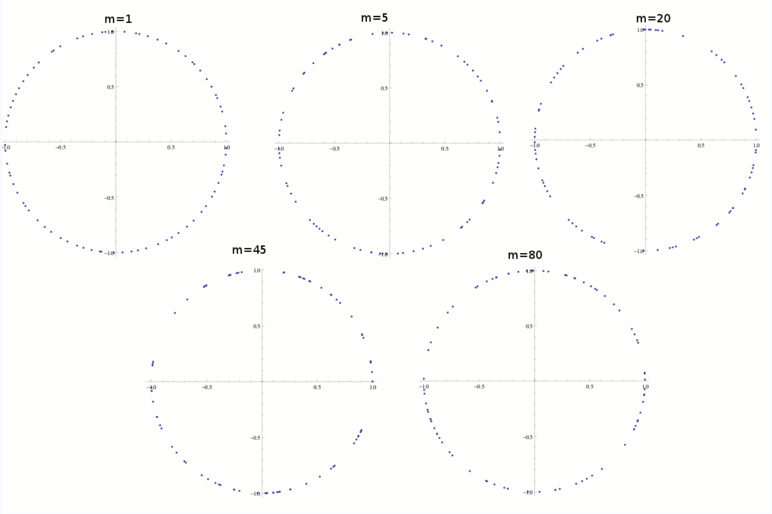
Let $U \in \mathbb{U}(n)$ be a random unitary matrix, and let $m \geq n$. Then the eigenvalues of U^m are distributed exactly *as n i.i.d. uniform points on $\mathbb{S}^1$.*

Theorem (Rains 2003)

Let $m \leq N$ be fixed. Then

$$[\mathbb{U}(N)]^m \overset{e.v.d.}{=} \bigoplus_{0 \leq j < m} \mathbb{U}\left(\left\lceil \frac{N-j}{m} \right\rceil \right),$$

where $\overset{e.v.d.}{=}$ denotes equality of eigenvalue distributions.



The eigenvalues of U^m for $m = 1, 5, 20, 45, 80$, for U a realization of a random 80×80 unitary matrix.

Theorem (E.M./M. Meckes)

Let ν denote the uniform probability measure on the circle and

$$W_p(\mu,\nu) := \inf\left\{ \left(\int |x-y|^p \, d\pi(x,y)\right)^{\frac{1}{p}} \;\middle|\; \begin{array}{l} \pi(A\times\mathbb{C}) = \mu(A) \\ \pi(\mathbb{C}\times A) = \nu(A) \end{array} \right\}.$$

Theorem (E.M./M. Meckes)

Let ν *denote the uniform probability measure on the circle and*

$$W_p(\mu,\nu) := \inf\left\{ \left(\int |x-y|^p \, d\pi(x,y)\right)^{\frac{1}{p}} \;\middle|\; \begin{array}{l} \pi(A\times\mathbb{C}) = \mu(A) \\ \pi(\mathbb{C}\times A) = \nu(A) \end{array} \right\}.$$

Then

- $\mathbb{E}\left[W_p(\mu_{m,N},\nu)\right] \le \frac{Cp\sqrt{m\left[\log\left(\frac{N}{m}\right)+1\right]}}{N}$.

Theorem (E.M./M. Meckes)

Let ν denote the uniform probability measure on the circle and

$$W_p(\mu,\nu) := \inf\left\{ \left(\int |x-y|^p \, d\pi(x,y)\right)^{\frac{1}{p}} \;\middle|\; \begin{array}{l} \pi(A\times\mathbb{C}) = \mu(A) \\ \pi(\mathbb{C}\times A) = \nu(A) \end{array} \right\}.$$

Then

- $\mathbb{E}\left[W_p(\mu_{m,N},\nu)\right] \leq \frac{Cp\sqrt{m\left[\log\left(\frac{N}{m}\right)+1\right]}}{N}.$
- *For* $1 \leq p \leq 2$,
$$\mathbb{P}\left[W_p(\mu_{m,N},\nu) \geq \frac{C\sqrt{m\left[\log\left(\frac{N}{m}\right)+1\right]}}{N} + t\right] \leq \exp\left[-\frac{N^2t^2}{24m}\right].$$
- *For* $p > 2$,
$$\mathbb{P}\left[W_p(\mu_{m,N},\nu) \geq \frac{Cp\sqrt{m\left[\log\left(\frac{N}{m}\right)+1\right]}}{N} + t\right] \leq \exp\left[-\frac{N^{1+\frac{2}{p}}t^2}{24m}\right].$$

Almost sure convergence

Corollary

For each N, let U_N *be distributed according to uniform measure on* $\mathbb{U}(N)$ *and let* $m_N \in \{1, \ldots, N\}$*. There is a C such that, with probability 1,*

$$W_p(\mu_{m_N,N}, \nu) \leq \frac{Cp\sqrt{m_N \log(N)}}{N^{\frac{1}{2}+\frac{1}{\max(2,p)}}}$$

eventually.

A miraculous representation of the eigenvalue counting function

A miraculous representation of the eigenvalue counting function

Fact: The set $\{e^{i\theta_j}\}_{j=1}^{N}$ of eigenvalues of U (uniform in $\mathbb{U}(N)$) is a determinantal point process.

A miraculous representation of the eigenvalue counting function

Fact: The set $\{e^{i\theta_j}\}_{j=1}^N$ of eigenvalues of U (uniform in $\mathbb{U}(N)$) is a determinantal point process.

Theorem (Hough/Krishnapur/Peres/Virág 2006)

Let $\mathcal{X}$ be a determinantal point process in Λ satisfying some niceness conditions. For $D \subseteq \Lambda$, let $\mathcal{N}_D$ be the number of points of $\mathcal{X}$ in D. Then

$$\mathcal{N}_D \stackrel{d}{=} \sum_k \xi_k,$$

where $\{\xi_k\}$ are independent Bernoulli random variables with means given explicitly in terms of the kernel of $\mathcal{X}$.

A miraculous representation of the eigenvalue counting function

That is, if $\mathcal{N}_\theta$ is the number of eigenangles of U between 0 and θ, then

$$\mathcal{N}_\theta \stackrel{d}{=} \sum_{j=1}^{N} \xi_j$$

for a collection $\{\xi_j\}_{j=1}^{N}$ of independent Bernoulli random variables.

A miraculous representation of the eigenvalue counting function

Recall Rains' second theorem:

$$[\mathbb{U}(N)]^m \overset{e.v.d.}{=} \bigoplus_{0 \le j < m} \mathbb{U}\left(\left\lceil \frac{N-j}{m} \right\rceil\right),$$

A miraculous representation of the eigenvalue counting function

Recall Rains' second theorem:

$$[\mathbb{U}(N)]^m \stackrel{e.v.d.}{=} \bigoplus_{0\le j<m} \mathbb{U}\left(\left\lceil \frac{N-j}{m} \right\rceil\right),$$

So: if $\mathcal{N}_{m,N}(\theta)$ denotes the number of eigenangles of U^m in $[0,\theta)$, then

$$\mathcal{N}_{m,N}(\theta) \stackrel{d}{=} \sum_{j=1}^{N} \xi_j,$$

for $\{\xi_j\}_{j=1}^N$ independent Bernoulli random variables.

Consequences of the miracle

Consequences of the miracle

- From Bernstein's inequality and the representation of $\mathcal{N}_{m,N}(\theta)$ as $\sum_{j=1}^{N} \xi_j$,

$$\mathbb{P}\left[\left|\mathcal{N}_{m,N}(\theta) - \mathbb{E}\mathcal{N}_{m,N}(\theta)\right| > t\right] \leq 2\exp\left[-\min\left\{\frac{t^2}{4\sigma^2}, \frac{t}{2}\right\}\right],$$

where $\sigma^2 = \operatorname{Var}\mathcal{N}_{m,N}(\theta)$.

Consequences of the miracle

- From Bernstein's inequality and the representation of $\mathcal{N}_{m,N}(\theta)$ as $\sum_{j=1}^{N} \xi_j$,

$$\mathbb{P}\left[\left|\mathcal{N}_{m,N}(\theta) - \mathbb{E}\mathcal{N}_{m,N}(\theta)\right| > t\right] \leq 2\exp\left[-\min\left\{\frac{t^2}{4\sigma^2}, \frac{t}{2}\right\}\right],$$

  where $\sigma^2 = \mathrm{Var}\,\mathcal{N}_{m,N}(\theta)$.

- $\mathbb{E}\mathcal{N}_{m,N}(\theta) = \frac{N\theta}{2\pi}$ (by rotation invariance).

Consequences of the miracle

- From Bernstein's inequality and the representation of $\mathcal{N}_{m,N}(\theta)$ as $\sum_{j=1}^{N} \xi_j$,

$$\mathbb{P}\left[\left|\mathcal{N}_{m,N}(\theta) - \mathbb{E}\mathcal{N}_{m,N}(\theta)\right| > t\right] \leq 2\exp\left[-\min\left\{\frac{t^2}{4\sigma^2}, \frac{t}{2}\right\}\right],$$

  where $\sigma^2 = \operatorname{Var}\mathcal{N}_{m,N}(\theta)$.

- $\mathbb{E}\mathcal{N}_{m,N}(\theta) = \frac{N\theta}{2\pi}$ (by rotation invariance).

- $\operatorname{Var}\left[\mathcal{N}_{1,N}(\theta)\right] \leq \log(N) + 1$ (e.g., via explicit computation with the kernel of the determinantal point process), and so

$$\operatorname{Var}\left(\mathcal{N}_{m,N}(\theta)\right) = \sum_{0 \leq j < m} \operatorname{Var}\left(\mathcal{N}_{1,\left\lceil \frac{N-j}{m} \right\rceil}(\theta)\right) \leq m\left(\log\left(\frac{N}{m}\right) + 1\right).$$

The concentration of $\mathcal{N}_{m,N}$ leads to concentration of individual eigenvalues about their predicted values:

$$\mathbb{P}\left[\left|\theta_j - \frac{2\pi j}{N}\right| > \frac{4\pi t}{N}\right] \leq 4\exp\left[-\min\left\{\frac{t^2}{m\left(\log\left(\frac{N}{m}\right)+1\right)}, t\right\}\right],$$

for each $j \in \{1, \ldots, N\}$:

The concentration of $\mathcal{N}_{m,N}$ leads to concentration of individual eigenvalues about their predicted values:

$$\mathbb{P}\left[\left|\theta_j - \frac{2\pi j}{N}\right| > \frac{4\pi t}{N}\right] \le 4\exp\left[-\min\left\{\frac{t^2}{m\left(\log\left(\frac{N}{m}\right)+1\right)}, t\right\}\right],$$

for each $j \in \{1, \ldots, N\}$:

$$\mathbb{P}\left[\theta_j > \frac{2\pi j}{N} + \frac{4\pi}{N}u\right] = \mathbb{P}\left[\mathcal{N}^{(m)}_{\frac{2\pi(j+2u)}{N}} < j\right]$$

The concentration of $\mathcal{N}_{m,N}$ leads to concentration of individual eigenvalues about their predicted values:

$$\mathbb{P}\left[\left|\theta_j - \frac{2\pi j}{N}\right| > \frac{4\pi t}{N}\right] \leq 4\exp\left[-\min\left\{\frac{t^2}{m\left(\log\left(\frac{N}{m}\right)+1\right)}, t\right\}\right],$$

for each $j \in \{1, \ldots, N\}$:

$$\mathbb{P}\left[\theta_j > \frac{2\pi j}{N} + \frac{4\pi}{N}u\right] = \mathbb{P}\left[\mathcal{N}^{(m)}_{\frac{2\pi(j+2u)}{N}} < j\right]$$
$$= \mathbb{P}\left[j + 2u - \mathcal{N}^{(m)}_{\frac{2\pi(j+2u)}{N}} > 2u\right]$$

The concentration of $\mathcal{N}_{m,N}$ leads to concentration of individual eigenvalues about their predicted values:

$$\mathbb{P}\left[\left|\theta_j - \frac{2\pi j}{N}\right| > \frac{4\pi t}{N}\right] \le 4\exp\left[-\min\left\{\frac{t^2}{m\left(\log\left(\frac{N}{m}\right)+1\right)}, t\right\}\right],$$

for each $j \in \{1, \ldots, N\}$:

$$\begin{aligned}\mathbb{P}\left[\theta_j > \frac{2\pi j}{N} + \frac{4\pi}{N}u\right] &= \mathbb{P}\left[\mathcal{N}^{(m)}_{\frac{2\pi(j+2u)}{N}} < j\right] \\ &= \mathbb{P}\left[j + 2u - \mathcal{N}^{(m)}_{\frac{2\pi(j+2u)}{N}} > 2u\right] \\ &\le \mathbb{P}\left[\left|\mathcal{N}^{(m)}_{\frac{2\pi(j+2u)}{N}} - \mathbb{E}\mathcal{N}^{(m)}_{\frac{2\pi(j+2u)}{N}}\right| > 2u\right].\end{aligned}$$

Bounding $\mathbb{E} W_p(\mu_{m,N}, \nu)$

If $\nu_N := \frac{1}{N} \sum_{j=1}^{N} \delta_{\exp\left(i \frac{2\pi j}{N}\right)}$, then $W_p(\nu_N, \nu) \leq \frac{\pi}{N}$ and

Bounding $\mathbb{E} W_p(\mu_{m,N}, \nu)$

If $\nu_N := \frac{1}{N} \sum_{j=1}^{N} \delta_{\exp\left(i \frac{2\pi j}{N}\right)}$, then $W_p(\nu_N, \nu) \leq \frac{\pi}{N}$ and

$$\mathbb{E} W_p^p(\mu_{m,N}, \nu_N) \leq \frac{1}{N} \sum_{j=1}^{N} \mathbb{E} \left| \theta_j - \frac{2\pi j}{N} \right|^p$$

Bounding $\mathbb{E} W_p(\mu_{m,N}, \nu)$

If $\nu_N := \frac{1}{N} \sum_{j=1}^{N} \delta_{\exp\left(i \frac{2\pi j}{N}\right)}$, then $W_p(\nu_N, \nu) \leq \frac{\pi}{N}$ and

$$\mathbb{E} W_p^p(\mu_{m,N}, \nu_N) \leq \frac{1}{N} \sum_{j=1}^{N} \mathbb{E} \left| \theta_j - \frac{2\pi j}{N} \right|^p$$

$$\leq 8\Gamma(p+1) \left(\frac{4\pi \sqrt{m \left[\log\left(\frac{N}{m}\right) + 1 \right]}}{N} \right)^p,$$

using the concentration result and Fubini's theorem.

Concentration of $W_p(\mu_{m,N}, \nu)$

Concentration of $W_p(\mu_{m,N}, \nu)$

The Idea: Consider the function $F_p(U) = W_p\left(\mu_{U^m}, \nu\right)$, where μ_{U^m} is the empirical spectral measure of U^m.

Concentration of $W_p(\mu_{m,N}, \nu)$

The Idea: Consider the function $F_p(U) = W_p\left(\mu_{U^m}, \nu\right)$, where μ_{U^m} is the empirical spectral measure of U^m.

- By Rains' theorem, it is distributionally the same as $F_p(U_1, \ldots, U_m) = \left(\frac{1}{m} \sum_{j=1}^{m} \mu_{U_j}, \nu\right)$.

Concentration of $W_p(\mu_{m,N}, \nu)$

The Idea: Consider the function $F_p(U) = W_p\left(\mu_{U^m}, \nu\right)$, where μ_{U^m} is the empirical spectral measure of U^m.

- By Rains' theorem, it is distributionally the same as $F_p(U_1, \ldots, U_m) = \left(\frac{1}{m}\sum_{j=1}^{m} \mu_{U_j}, \nu\right)$.
- $F_p(U_1, \ldots, U_m)$ is Lipschitz (w.r.t. the L_2 sum of the Euclidean metrics) with Lipschitz constant $N^{-\frac{1}{\max(p,2)}}$.

Concentration of $W_p(\mu_{m,N}, \nu)$

The Idea: Consider the function $F_p(U) = W_p\left(\mu_{U^m}, \nu\right)$, where μ_{U^m} is the empirical spectral measure of U^m.

- By Rains' theorem, it is distributionally the same as $F_p(U_1, \ldots, U_m) = \left(\frac{1}{m}\sum_{j=1}^{m} \mu_{U_j}, \nu\right)$.
- $F_p(U_1, \ldots, U_m)$ is Lipschitz (w.r.t. the L_2 sum of the Euclidean metrics) with Lipschitz constant $N^{-\frac{1}{\max(p,2)}}$.
- If we had a general concentration phenomenon on $\bigoplus_{0\le j<m} \mathbb{U}\left(\left\lceil \frac{N-j}{m} \right\rceil\right)$, concentration of $W_p\left(\mu_{U^m}, \nu\right)$ would follow.

Concentration on $\mathbb{U}(N_1) \oplus \cdots \oplus \mathbb{U}(N_k)$

Theorem (E. M./M. Meckes)

Given $N_1, \ldots, N_k \in \mathbb{N}$*, denote by* $M = \mathbb{U}(N_1) \times \cdots \mathbb{U}(N_k)$ *equipped with the* L_2*-sum of Hilbert–Schmidt metrics.*

Suppose that $F : M \to \mathbb{R}$ *is L-Lipschitz, and that* $U_j \in \mathbb{U}(N_j)$ *are independent, uniform random unitary matrices, for* $1 \leq j \leq k$*. Then for each* $t > 0$*,*

$$\mathbb{P}\Big[F(U_1, \ldots, U_k) \geq \mathbb{E}F(U_1, \ldots, U_k) + t\Big] \leq e^{-Nt^2/12L^2},$$

where $N = \min\{N_1, \ldots, N_k\}$*.*