Random matrix theory: lecture 2

Finite-size analysis (part I)

Problem: let $H$ be a $n \times n$ random matrix with a given distribution; what can be said about the joint distribution of its eigenvalues $p(\lambda_1, \ldots, \lambda_n)$?

Linear algebra reminder ($H$ $n \times n$ complex matrix)

- $H$ is said to be diagonalizable if there exist an invertible matrix $S$ and a diagonal matrix $\Lambda$ such that $H = S \Lambda S^{-1}$

  In this case, $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$ where $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of $H$.

- Not every matrix $H$ is diagonalizable, but the following is true in general: there always exist an invertible matrix $S$ and an upper triangular matrix $T$ such that $H = STS^{-1}$

  Moreover, $T$ is block-diagonal, with blocks of the form $\begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$ [Jordan decomposition]

  Again, the elements of the diagonal of $T$ are the eigenvalues of $H$. 
Particular cases:

- If $H$ is normal, i.e. $HH^* = H^*H$, then $H$ is unitarily diagonalizable, i.e. there exist a unitary matrix $U$ (i.e. $UU^* = I$) and a diagonal matrix $\Lambda$ such that $H = U \Lambda U^*$

**NB:** This is known as the spectral theorem

- There are three important sub-cases of the above:
  a) If $H$ is Hermitian, i.e. $H = H^*$, then $H$ is normal and $H = U \Lambda U^*$, where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$ and the eigenvalues $\lambda_1, \ldots, \lambda_n$ are real
  b) If $H$ is non-negative definite, i.e. $x^*Hx \geq 0$ for any vector $x \in \mathbb{C}^n$, then $H$ is normal and $H = U \Lambda U^*$, where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$ and the eigenvalues $\lambda_1, \ldots, \lambda_n$ are non-negative
  c) If $H$ is unitary, i.e. $HH^* = I$, then $H$ is normal and $H = U \Lambda U^*$, where $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$ and the eigenvalues $\lambda_1, \ldots, \lambda_n$ are of modulus 1 (i.e. $|\lambda_j| = 1 \forall j$)

\(*\) $H$ is Hermitian, so...
For reasons that will become apparent in the class, it is (much) easier to deal with random matrices whose eigenvalues are distributed on a particular curve in the complex plane, and not in the whole plane. We will therefore focus on the last three subcases.

Back to the joint eigenvalue distribution problem.

General strategy: given an ensemble of normal random matrices $H$, we may interpret the spectral theorem $H = U \Lambda U^*$ as a change of variables $H \mapsto (\Lambda, U)$.

Provided that $H$ is distributed according to $p(H)$, we therefore have $p(H) \, dH = p(U \Lambda U^*) \, |J(\Lambda, U)| \, d\Lambda \, dU,$

where $J(\Lambda, U)$ is the Jacobian of the change of variables. The joint distribution of $(\Lambda, U)$ is given by

$$p(\Lambda, U) = p(U \Lambda U^*) \cdot |J(\Lambda, U)|$$

eigenvalues $\leftrightarrow$ eigenvectors

And as we will see, this expression simplifies drastically in some particular cases.
Warm-up (case n=1!!)

Let \( x, y \) be iid r.v. \( \sim N_{iR}(0, \frac{1}{2}) \), i.e. their joint density is given by

\[
p(x, y) = \frac{1}{\sqrt{\pi}} \exp(-x^2) \cdot \frac{1}{\sqrt{\pi}} \exp(-y^2) = \frac{1}{\pi} e^{-x^2 - y^2}
\]

**NB:** the complex r.v. \( z = x + iy \) has therefore a density

\[
p(z) = \frac{1}{\pi} e^{-|z|^2}; \text{ notation: } z \sim N_{iC}(0, 1)
\]

Let us consider the change of variable \( x + iy = re^{i\theta} \).

The Jacobian is given by

\[
J(r, \theta) = \det \left( \begin{array}{cc} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{array} \right)
\]

\[
= \det \left( \begin{array}{cc} \cos \theta & \sin \theta \\ -r \sin \theta & r \cos \theta \end{array} \right) = r
\]

Therefore,

\[
p(r, \theta) = p(x(r, \theta), y(r, \theta)) \cdot r = \frac{1}{\pi} e^{-r^2} \cdot r
\]

\[
= 2r e^{-r^2} \cdot \frac{1}{2\pi} \cdot \frac{1}{\sqrt{\pi}} 
\]

(Rayleigh dist.) \( \tilde{p}(r) \) does actually not depend on \( \theta \); this implies:

a) the distribution is uniform in \( \theta \)

b) \( r \) and \( \theta \) are independent (since factorization)

c) for any given \( \theta \), \( z \) and \( ze^{i\theta} \) have the same distribution, deterministic

\(*\) In addition, since \( p(z) \) depends only on \( |z| \),

The r.v. \( z \) is said to be "circularly symmetric"
**Gaussian Orthogonal Ensemble (GOE)**

Let $H$ be a $n \times n$ real symmetric random matrix such that:

- $\{h_{jk}, j \leq k\}$ are independent r.v. ($h_{jk} \sim N_R(0,1)$)
- $h_{jj} \sim N_R(0,1)$, $h_{jk} \sim N_R(0, \frac{1}{2})$ for $j < k$

**Distribution of $H$:**

$$p(H) = \prod_{j=1}^{n} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{h_{jj}^2}{2}\right) \cdot \prod_{j<k} \frac{1}{\sqrt{\pi}} \exp\left(-\frac{h_{jk}^2}{2}\right)$$

$$= C_n \exp\left(-\frac{1}{2} \sum_{j=1}^{n} h_{jj}^2 + \sum_{j<k} h_{jk}^2\right)$$

$$= C_n \exp\left(-\frac{1}{2} \sum_{j=1}^{n} h_{jj}^2 + \frac{1}{2} \sum_{j<k} h_{jk}^2\right)$$

$$= C_n \exp\left(-\frac{1}{2} \text{Tr}(H^2)\right)$$

since $H = H^T$

By the spectral theorem, there exists an $n \times n$ **orthogonal** matrix $V$ (i.e. $VV^T = I$) and $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$, with $\lambda_1, \ldots, \lambda_n$ real, such that $H = V \Lambda V^T$

i.e. $h_{jk} = \sum_{\ell=1}^{n} \lambda_{\ell} v_{j\ell} v_{k\ell}$ for $j, k = 1, \ldots, n$

- **Sanity check:** how many free (real) parameters do we have on each side?

  - **On the left:** $n$ diag. parameters $+ \frac{n(n-1)}{2}$ upper diag. parameters

  $$= \frac{n(n+1)}{2} \text{ parameters}$$
on the right: \( n \) diag. parameters in \( \Lambda \)

\((\Lambda, V)\) + \(\frac{n(n-1)}{2}\) parameters in \( V \) (see construction below)

= \(\frac{n^2-n}{2}\) parameters \( V \)

Aside: how many free parameters are there

in an orthogonal matrix \( V \)?

Reminder: \( V V^T = I \) means the rows of \( V \) are orthonormal

Vectors \( v_1, \ldots, v_n \) in \( \mathbb{R}^n \)

So:

• for the first row \( v_1 \), there are \( n-1 \) free parameters
  (since \( v_1 v_1^T = I \))

• for the second row \( v_2 \), there are \( n-2 \) free parameters
  (since \( v_2 v_2^T = I \) & \( v_2^T v_1 = 0 \))

• etc.

In total, this leads to \( (n-1) + (n-2) + \ldots + 1 + 0 = \frac{n(n-1)}{2} \) parameters.

Jacobian:

The Jacobian of the change of variables \( \mathbf{H} \rightarrow (\Lambda, V) \)

is given by:

\[ \begin{vmatrix} \frac{\partial \Lambda}{\partial \lambda_j} & \frac{\partial \Lambda}{\partial \lambda_k} \\ \frac{\partial V}{\partial \lambda_j} & \frac{\partial V}{\partial \lambda_k} \end{vmatrix} \]

Result of the computation:

\[ \begin{vmatrix} \frac{\partial \Lambda}{\partial \lambda_j} & \frac{\partial \Lambda}{\partial \lambda_k} \\ \frac{\partial V}{\partial \lambda_j} & \frac{\partial V}{\partial \lambda_k} \end{vmatrix} = \mathbf{H} \begin{pmatrix} \Lambda & -I \end{pmatrix} \]

[ Homework 1: explicit simple case \((n=2)\)]
Heuristics for the above computation:

- \( h_{jik} = \sum_{e=1}^{n} \lambda_e V_{je} V_{ke} \)

\[ \Rightarrow \begin{cases} \frac{\partial h_{jik}}{\partial V_{em}} = V_{je} V_{ke} = \text{cst} \quad \forall \ i \neq j \neq k \\ \frac{\partial h_{jik}}{\partial V_{em}} = (\delta_{je} V_{km} + \delta_{ke} V_{jm}) \lambda_e = \text{linear fn in } \lambda_e \end{cases} \]

\[ \Rightarrow \{ \frac{\partial}{\partial \lambda_e} \frac{\partial h_{jik}}{\partial V_{em}} = \frac{\partial^2 h_{jik}}{\partial V_{em} \partial V_{pm}} \} \quad \text{i.e. } \beta = 0 \]

so the only polynomial satisfying these two conditions is of the form \( \prod_{j<k} (\lambda_j - \lambda_k) \)

**NB**: a remarkable fact is that \( \beta \) does not depend on \( V \) (similar to the polar coordinates example)!

Conclusion for the GAE:

\( \tilde{p}(\Lambda, V) = p(V \Lambda V^T) \prod_{j<k} (\lambda_j - \lambda_k) \)

\[
= C_n \exp \left( -\frac{1}{2} \text{Tr} \left( (V \Lambda V^T)^2 \right) \right) \prod_{j<k} (\lambda_j - \lambda_k) \\
= \text{Tr} (V \Lambda V^T) \\
= \text{Tr} (V \Lambda^2 V^T) = \text{Tr} (\Lambda^2 V^T V) \\
= \text{Tr} (\Lambda^2) \\
= C_n \exp \left( -\frac{1}{2} \sum_{j=1}^{n} \lambda_j^2 \right) \prod_{j<k} (\lambda_j - \lambda_k) 
\]
Same remark as before:

\( \tilde{\rho}(\Lambda, V) \) does not depend on \( V \) at all.

\( \Rightarrow \)

a) the distribution of \( V \) is uniform over the set of orthogonal matrices ("Haar distribution")

b) \( \Lambda \) and \( V \) are actually independent, i.e.

the eigenvalues of \( H \) are independent from its eigenvectors!

c) for any given deterministic orthogonal matrix \( W \), one obtains that \( H \) and \( WTHW \)
have the same distribution, i.e. the distribution of \( H \) is invariant under orthogonal transformations, therefore the name of the ensemble.

NB: the above computation was made possible by the fact that the distribution of \( H \) only depends on \( Tr(\Lambda^2) = Tr(\Lambda^2) \).