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 $p(\lambda_1 .. \lambda_n)$?

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 Λ such that $H = S \Lambda S^{-1}$

× in this case, $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_n)$ where $\lambda_1, \dots, \lambda_n$ 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 $H = S T S^{-1}$

moreover, T is block-diagonal, with blocks

× of the form $\begin{pmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$ [Jordan decomposition]

again, the elements of the diagonal of T are the eigenvalues of H

Particular cases:

[complex conjugate transpose]

- 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 Λ 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 $\Lambda = \text{diag}(\lambda_1 .. \lambda_n)$ and the eigenvalues $\lambda_1 .. \lambda_n$ are real

× b) if H is non-negative definite, i.e. $x^* H x \geq 0$ for any vector $x \in \mathbb{C}^n$, then(*) H is normal and $H = U\Lambda U^*$, where $\Lambda = \text{diag}(\lambda_1 .. \lambda_n)$ and the eigenvalues $\lambda_1 .. \lambda_n$ are non-negative

× c) if H is unitary, i.e. $HH^* = I$, then H is normal and $H = U\Lambda U^*$, where $\Lambda = \text{diag}(\lambda_1 .. \lambda_n)$ and the eigenvalues $\lambda_1 .. \lambda_n$ 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.

<u>Back to the joint eigenvalue distribution problem</u>

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

Provided that H is distributed according to $p(H)$,

× we therefore have $p(H)\, dH = p(U \Lambda U^*) \cdot |J(\Lambda, U)|\, d\Lambda\, dU$,

where $J(\Lambda, U)$ is the Jacobian of the change of variables.

The joint distribution of (Λ, U) is given by

×
$$\tilde{p}(\underset{\text{eigenvalues}}{\Lambda}, \underset{\text{eigenvectors}}{U}) = p(U \Lambda U^*) \cdot |J(\Lambda, U)|$$

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. $\sim N_{\mathbb{R}}(0, \frac{1}{2})$, 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 $p(z) = \frac{1}{\pi} e^{-|z|^2}$; notation: $z \sim N_{\mathbb{C}}(0,1)$ (*)

Let us consider the change of variable $x + iy = re^{i\theta}$:

$\left[\text{i.e. } \begin{matrix} x = r\cos\theta \\ y = r\sin\theta \end{matrix}\right]$

The Jacobian is given by

$$J(r,\theta) = \det\begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix}$$

$$= \det\begin{pmatrix} \cos\theta & \sin\theta \\ -r\sin\theta & r\cos\theta \end{pmatrix} = r$$

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

$$= \underbrace{2r\, e^{-r^2}}_{\text{(Rayleigh dist.) } \tilde{p}(r)} \cdot \underbrace{\frac{1}{2\pi}}_{\tilde{p}(\theta)}$$

Remark:

$\tilde{p}(r,\theta)$ does actually not depend on θ; this implies:

a) the distribution is uniform in θ

b) r and θ are independent (since factorization)

c) for any given φ, z and $ze^{i\varphi}$ have the same distribution
deterministic

(*) in addition, since $p(z)$ depends only on $|z|$, the r.v. z is said to be "circularly symmetric"

× name

Gaussian Orthogonal Ensemble (GOE)

Let H be a $n \times n$ real symmetric random matrix such that:
- $\{h_{jk},\ j \le k\}$ are independent r.v. ($\&\ h_{kj} = h_{jk}$)
- $h_{jj} \sim N_{\mathbb{R}}(0,1)$, $h_{jk} \sim N_{\mathbb{R}}(0,\frac{1}{2})$ $\forall j<k$

- <u>Distribution of H</u>:

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

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

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

$$= C_n \exp\left(-\frac{1}{2}\sum_{j,k} h_{jk}^2\right) = C_n \exp\left(-\frac{1}{2}\,\mathrm{Tr}(HH^T)\right)$$

$$= C_n \exp\left(-\frac{1}{2}\,\mathrm{Tr}(H^2)\right) \quad \text{since } H = H^T$$

real case

By the spectral theorem, there exists a $n\times n$ <u>orthogonal</u> matrix V (i.e. $VV^T = I$) and $\Lambda = \mathrm{diag}(\lambda_1 .. \lambda_n)$, with $\lambda_1 .. \lambda_n$ real, such that

$$\left.\begin{aligned} & H = V \Lambda V^T \\ & \text{i.e. } h_{jk} = \sum_{l=1}^{n} \lambda_l v_{jl} v_{kl} \qquad j,k = 1..n \end{aligned}\right\} \to \text{change of variables}$$

- <u>Sanity check</u>: how many free (real) parameters do we have on each side?

  - <u>on the left</u> (H): 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 Λ
  (Λ, V) $+ \frac{n(n-1)}{2}$ parameters in V (see construction below)
  $= \frac{n^2+n}{2}$ parameters ✓

Aside: how many free parameters are there in an orthogonal matrix V?

Reminder: $VV^T = I$ means the rows of V are orthonormal vectors $v_1, \dots, v_n$ in $\mathbb{R}^n$

so:
- for the first row v_1, there are $n-1$ free parameters (since $\|v_1\| = 1$)
- for the second row v_2, there are $n-2$ " (since $\|v_2\| = 1$ & $v_2 v_1^T = 0$)
- etc.
- in total, this leads to $(n-1) + (n-2) + \dots + 1 + 0 = \frac{n(n-1)}{2}$ param. ✓

- Jacobian:

The Jacobian of the change of variables $H \mapsto (\Lambda, V)$ is given by:
$$\mathcal{J}(\{\lambda_\ell\}, \{v_{\ell m}\}) = \det\left(\left\{\frac{\partial h_{jk}}{\partial \lambda_\ell}\right\} \Bigg| \left\{\frac{\partial h_{jk}}{\partial v_{\ell m}}\right\}\right)$$

$\frac{n(n+1)}{2} \times n$ matrix

$\frac{n(n+1)}{2} \times \frac{n(n-1)}{2}$ matrix

Result of the computation:
$$\mathcal{J}(\{\lambda_\ell\}, \{v_{\ell m}\}) = \prod_{j<k} (\lambda_j - \lambda_k)$$

[Homework 1 → explicit simple case $(n=2)$]

Heuristics for the above computation:

- $h_{jk} = \sum_{\ell=1}^{n} \lambda_\ell v_{j\ell} v_{k\ell}$

$$\Longrightarrow \begin{cases} \dfrac{\partial h_{jk}}{\partial \lambda_\ell} = v_{j\ell} v_{k\ell} = \text{cst in } \lambda \\ \dfrac{\partial h_{jk}}{\partial v_{\ell m}} = (\delta_{j\ell} v_{km} + \delta_{k\ell} v_{jm}) \lambda_\ell = \text{linear fn in } \lambda \end{cases}$$

$$\Longrightarrow \begin{cases} J = \det(\;\cdot\;) = \text{polynomial in } \lambda \text{ of max. degree } \frac{n(n-1)}{2} \\ \text{if } \lambda_\ell = \lambda_p, \text{ then } \dfrac{\partial h_{jk}}{\partial v_{\ell m}} = \dfrac{\partial h_{jk}}{\partial v_{pm}} \text{ i.e. } J = 0 \end{cases}$$

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 J does not depend on V (similar to the polar coordinates example)!

[to be proven]

- Conclusion for the GOE:

$$\tilde{p}(\Lambda, V) = p(V\Lambda V^T) \cdot |J(\Lambda, V)|$$

$$= C_n \exp\Big(-\frac{1}{2} \underbrace{\mathrm{Tr}((V\Lambda V^T)^2)}\Big) \cdot \prod_{j<k} |\lambda_j - \lambda_k|$$

$$= \mathrm{Tr}(V\Lambda V^T V \Lambda V^T)$$

$$= \mathrm{Tr}(V\Lambda^2 V^T) = \mathrm{Tr}(\Lambda^2 V^T V)$$

$$= \mathrm{Tr}(\Lambda^2)$$

$$= C_n \exp\Big(-\frac{1}{2}\sum_{j=1}^{n} \lambda_j^2\Big) \cdot \prod_{j<k} |\lambda_j - \lambda_k|$$

Same remark as before:

$\tilde{p}(\Lambda, V)$ does not depend on V at all

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

b) Λ 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 $W^T H W$ 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 $\mathrm{Tr}(H^2) = \mathrm{Tr}(\Lambda^2)$.