Random matrices and communication systems: WEEK 6

1 Wishart random matrices: joint eigenvalue distribution

1.1 Real case

Recall first from previous lecture that if $W = HH^T$, where H is an $n \times m$ random matrix with i.i.d.$\sim \mathcal{N}_{\mathbb{R}}(0,1)$ entries and $m \geq n$, then the joint distribution of the entries of W is given by

$$p_W(W) = c_{n,m} \det(W)^{\frac{m-n-1}{2}} \exp\left(-\frac{1}{2}\operatorname{Tr}(W)\right) 1_{\{W \geq 0\}}$$

By the spectral theorem, the matrix W is orthogonally diagonalizable, that is, there exist V an $n \times n$ orthogonal matrix (i.e. $VV^T = I$) and $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ such that $\lambda_j \geq 0$ for all $1 \leq j \leq n$ and $W = V\Lambda V^T$, i.e.

$$w_{jk} = \sum_{l=1}^{n} \lambda_l \, v_{jl} \, v_{kl}, \quad 1 \leq j, k \leq n$$

Again, this can be viewed as a change of variables; on the left-hand side, there are n diagonal free parameters w_{jj} and $\frac{n(n-1)}{2}$ off-diagonal free parameters w_{jk}, $j < k$ in the matrix W (the remaining off-diagonal parameters are fixed, as W is symmetric); on the right-hand side, there are n free parameters in the matrix Λ and $(n-1) + (n-2) + \ldots + 1 + 0 = \frac{n(n-1)}{2}$ free parameters in the matrix V. So the number of free parameters on both sides coincide. The joint distribution of Λ and V is then given by

$$p_{\Lambda,V}(\Lambda, V) = p_W\left(V\Lambda V^T\right) |J(\Lambda, V)|$$

where, according to the above formula,

$$\begin{aligned} p_W\left(V\Lambda V^T\right) &= c_{n,m} \det\left(V\Lambda V^T\right)^{\frac{m-n-1}{2}} \exp\left(-\frac{1}{2}\operatorname{Tr}\left(V\Lambda V^T\right)\right) \\ &= c_{n,m} \det(\Lambda)^{\frac{m-n-1}{2}} \exp\left(-\frac{1}{2}\operatorname{Tr}(\Lambda)\right) \end{aligned}$$

and $J(\Lambda, V)$ is the Jacobian of the transformation $W \mapsto (\Lambda, V)$. We now set out to compute this Jacobian. Let $N = \frac{n(n-1)}{2}$ and let us denote by $p_1, \ldots, p_N$ the N free parameters in the matrix V (leaving aside the explicit description of what these N parameters are: we will see in the following that this is actually not needed). The Jacobian is then given by

$$J(\Lambda, V) = \det \left(\begin{array}{c|c} \left\{\frac{\partial w_{jj}}{\partial \lambda_i}\right\}_{i,j} & \left\{\frac{\partial w_{jk}}{\partial \lambda_i}\right\}_{i,j<k} \\ \hline \left\{\frac{\partial w_{jj}}{\partial p_i}\right\}_{i,j} & \left\{\frac{\partial w_{jk}}{\partial p_i}\right\}_{i,j<k} \end{array}\right)$$

where the blocks in the above matrix are respectively of size $n \times n$, $n \times N$, $N \times n$ and $N \times N$. As $W = V\Lambda V^T$, the computation of the above partial derivatives gives, in matrix form:

$$\begin{aligned} \frac{\partial W}{\partial \lambda_i} &= V\Delta^{(i)}V^T, \quad \text{where } \Delta^{(i)}_{jk} = \delta_{ij}\,\delta_{ik} \\ \frac{\partial W}{\partial p_i} &= \frac{\partial V}{\partial p_i}\Lambda V^T + V\Lambda\frac{\partial V^T}{\partial p_i} \end{aligned}$$

Multiplying both these equations by V^T and V and the left-hand and right-hand side, we obtain

$$\begin{aligned} V^T\frac{\partial W}{\partial \lambda_i}V &= \Delta^{(i)} \\ V^T\frac{\partial W}{\partial p_i}V &= \left(V^T\frac{\partial V}{\partial p_i}\right)\Lambda + \Lambda\left(\frac{\partial V^T}{\partial p_i}V\right) \end{aligned}$$

Let now $S^{(i)} = V^T \frac{\partial V}{\partial p_i}$. As $V^T V = I$, we also obtain

$$V^T \frac{\partial V}{\partial p_i} + \frac{\partial V^T}{\partial p_i} V = 0, \quad \text{i.e.} \quad \frac{\partial V^T}{\partial p_i} V = -S^{(i)}$$

which allows us to rewrite $V^T \frac{\partial W}{\partial p_i} V = S^{(i)} \Lambda - \Lambda S^{(i)}$. Component-wise, the two equations for the derivatives with respect to λ_i and p_i therefore read:

$$\begin{cases} \displaystyle\sum_{l,m=1}^{n} \frac{\partial w_{lm}}{\partial \lambda_i} \, v_{lj} \, v_{mk} = \delta_{ij} \, \delta_{ik} \\ \displaystyle\sum_{l,m=1}^{n} \frac{\partial w_{lm}}{\partial p_i} \, v_{lj} \, v_{mk} = S_{jk}^{(i)} \, (\lambda_k - \lambda_j) \end{cases} \tag{1}$$

With the help of these formulas, let us now compute the Jacobian $J(\Lambda, V)$ when $V = I$. In this case, the above two formulas boil down to

$$\frac{\partial w_{jk}}{\partial \lambda_i} = \delta_{ij} \, \delta_{ik} \quad \text{and} \quad \frac{\partial w_{jk}}{\partial p_i} = S_{jk}^{(i)} \, (\lambda_k - \lambda_j)$$

so

$$\begin{aligned} J(\Lambda, V) &= \det \begin{pmatrix} I & 0 \\ 0 & \left\{ S_{jk}^{(i)} \, (\lambda_k - \lambda_j) \right\}_{i,j<k} \end{pmatrix} = \det \left(\left\{ S_{jk}^{(i)} \, (\lambda_k - \lambda_j) \right\}_{i,j<k} \right) \\ &= \prod_{j<k} (\lambda_k - \lambda_j) \det \left(\left\{ S_{jk}^{(i)} \right\}_{i,j<k} \right) = \prod_{j<k} (\lambda_k - \lambda_j) \, f(V) \end{aligned}$$

for some function f, as the matrix elements $S_{jk}^{(i)}$ possibly only depend on V. We claim that the same conclusion holds in the case where $V \neq I$. To this end, let us consider

$$\widetilde{J}(\Lambda, V) = \det \left(\begin{pmatrix} \left\{ \frac{\partial w_{ll}}{\partial \lambda_i} \right\}_{i,l} & \left\{ \frac{\partial w_{lm}}{\partial \lambda_i} \right\}_{i,l<m} \\ \left\{ \frac{\partial w_{ll}}{\partial p_i} \right\}_{i,l} & \left\{ \frac{\partial w_{jk}}{\partial p_i} \right\}_{i,l<m} \end{pmatrix} \begin{pmatrix} \left\{ v_{lj}^2 \right\}_{l,j} & \left\{ v_{lj} v_{lk} \right\}_{l,j<k} \\ \left\{ 2 v_{lj} v_{mj} \right\}_{l<m,j} & \left\{ 2 v_{lj} v_{mk} \right\}_{l<m,j<k} \end{pmatrix} \right)$$

Using the fact that $\det(AB) = \det(A) \det(B)$ and observing that the second term on the right-hand side only depends on V, we deduce that $\widetilde{J}(\Lambda, V) = J(\Lambda, V) \, g(V)$ for some function g. On the other hand, performing the matrix multiplication inside the determinant gives for matrix element $i, (jk)$ in the first n rows:

$$\sum_{l=1}^{n} \frac{\partial w_{ll}}{\partial \lambda_i} \, v_{lj} \, v_{lk} + 2 \sum_{l<m} \frac{\partial w_{lm}}{\partial \lambda_i} \, v_{lj} \, v_{mk} = \sum_{l,m=1}^{n} \frac{\partial w_{lm}}{\partial \lambda_i} \, v_{lj} \, v_{mk}$$

and likewise for the matrix element $i, (jk)$ in the last N rows:

$$\sum_{l=1}^{n} \frac{\partial w_{ll}}{\partial p_i} \, v_{lj} \, v_{lk} + 2 \sum_{l<m} \frac{\partial w_{lm}}{\partial p_i} \, v_{lj} \, v_{mk} = \sum_{l,m=1}^{n} \frac{\partial w_{lm}}{\partial p_i} \, v_{lj} \, v_{mk}$$

Using then equation (1), we obtain again

$$\widetilde{J}(\Lambda, V) = \det \begin{pmatrix} I & 0 \\ 0 & \left\{ S_{jk}^{(i)} \, (\lambda_k - \lambda_j) \right\}_{i,j<k} \end{pmatrix} = \prod_{j<k} (\lambda_k - \lambda_j) \, f(V)$$

which, together with the above observation that $\widetilde{J}(\Lambda, V) = J(\Lambda, V) \, g(V)$, proves the claim. What can be deduced so far from all these computations is that

$$p_{\Lambda,V}(\Lambda, V) = c_{n,m} \det(\Lambda)^{\frac{m-n-1}{2}} \exp\left(-\frac{1}{2} \operatorname{Tr}(\Lambda) \right) \prod_{j<k} |\lambda_k - \lambda_j| \, |f(V)|$$

for some function f. This is actually saying that $p_{\Lambda,V}(\Lambda, V) = p_\Lambda(\Lambda)\, p_V(V)$, so the eigenvalues and eigenvectors of W are independent! The joint distribution of the eigenvalues is given by

$$p(\lambda_1, \ldots, \lambda_n) = c_{n,m} \prod_{j=1}^{n} \left(\lambda_j^{\frac{m-n-1}{2}} \exp(-\lambda_j/2)\, 1_{\lambda_j \geq 0} \right) \prod_{j<k} |\lambda_k - \lambda_j|$$

where $c_{n,m}$ is the normalization constant, which can be computed explicitly; it differs from the constant in the expression for $p_W(W)$, but in order to keep notation simple, we do not change notation here.

The above distribution may also be rewritten in the following form:

$$p(\lambda_1, \ldots, \lambda_n) = c_{n,m} \exp\left(-\sum_{j=1}^{n} \left(\frac{\lambda_j}{2} - \frac{m-n-1}{2} \log(\lambda_j) \right) + \sum_{j<k} \log|\lambda_k - \lambda_j| \right) 1_{\lambda_1 \geq 0, \ldots, \lambda_n \geq 0}$$

and given the following interpretation: it represents the Gibbs distribution of a system of n particles in positions $\lambda_1, \ldots, \lambda_n$ evolving in a potential $U(\lambda) = \frac{\lambda}{2} - \frac{m-n-1}{2} \log(\lambda)$ and repelling each other. Two opposite forces operate here: on one hand, the particles would all like to be in the minimum of the potential, but as they repel each other, there is not enough room for them, so some are driven away from this minimum. The fact that eigenvalues repel each other is a common feature to most random matrix models (essentially because of the term resulting from the above Jacobian computation).

Now, what is the distribution of the eigenvectors? As already seen, for every fixed $n \times n$ orthogonal matrix $O \in O(n)$, the matrices H and OH share the same distribution. Therefore, so do the matrices W and OWO^T, which is saying that

$$V \Lambda V^T \quad \text{and} \quad (OV)\Lambda(OV)^T$$

also share the same distribution. By the independence of Λ and V (and the non-singularity of the distribution of Λ), this finally implies that V and OV share the same distribution, for every fixed $n \times n$ orthogonal matrix O. This in turn implies that the matrix V is distributed according to the Haar distribution on $O(n)$, which is the unique distribution on $O(n)$ being invariant under orthogonal transformations.

1.2 Complex case

We shall not repeat here the whole reasoning; we just mention the main steps of the computation. In this case, $W = HH^*$, where H is an $n \times m$ random matrix with i.i.d.$\sim \mathcal{N}_\mathbb{C}(0,1)$ entries and $m \geq n$. The joint distribution of the entries of W is given by

$$p_W(W) = c_{n,m} \det(W)^{m-n} \exp(-\mathrm{Tr}(W))\, 1_{\{W \geq 0\}}$$

By the spectral theorem, the matrix W is unitarily diagonalizable, that is, there exist U an $n \times n$ unitary matrix (i.e. $UU^* = I$) and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ such that $\lambda_j \geq 0$ for all $1 \leq j \leq n$ and $W = U\Lambda U^*$, i.e.

$$w_{jk} = \sum_{l=1}^{n} \lambda_l\, u_{jl}\, \overline{u_{kl}}, \quad 1 \leq j, k \leq n$$

On the left-hand side, there are n diagonal free real parameters and $2\,\frac{n(n-1)}{2} = n^2 - n$ off-diagonal free real parameters in the matrix W; on the right-hand side, there are n free real parameters in the matrix Λ and n^2 free real parameters in the matrix U. So here we see that there is a mismatch. The mismatch can be resolved by observing that in the complex case, an eigenvector rotated by $e^{i\phi}$ remains an eigenvector. So it is possible to set the first component of each eigenvector of W to be a real number, which reduces by n the numbers of free real parameters in U, so that the number of free real parameters on both sides coincide. The joint distribution of Λ and U is then given by

$$p_{\Lambda,U}(\Lambda, U) = p_W(U\Lambda U^*)\, |J(\Lambda, U)|$$

where

$$p_W\left(U\Lambda U^*\right) = c_{n,m}\det(\Lambda)^{m-n}\,\exp(-\mathrm{Tr}(\Lambda))$$

and the computation of the Jacobian gives

$$J(\Lambda, U) = \prod_{j<k}(\lambda_k - \lambda_j)^2\, f(U)$$

for some function f. Therefore, Λ and U are also independent in this case, and

$$p(\lambda_1, \dots, \lambda_n) = c_{n,m}\prod_{j=1}^{n}\left(\lambda_j^{m-n}\,\exp(-\lambda_j)\,1_{\lambda_j\geq 0}\right)\prod_{j<k}(\lambda_k - \lambda_j)^2$$

where $c_{n,m}$ is the normalization constant, which differs from the previous $c_{n,m}$ and can be computed explicitly. Finally, similarly to the previous section, U is distributed according to the Haar distribution on $U(n)$, which is the unique distribution on $U(n)$ being invariant under unitary transformations.