Uniqueness of the Gaussian Orthogonal Ensemble

José Ángel Sánchez Gómez,† Victor Amaya Carvajal!

December, 2017.

Abstract

It is not hard to prove that, given a random Wigner matrix \( X \) that belongs to the Gaussian Orthogonal Ensemble (GOE), then \( X \) is invariant under orthogonal conjugations. Our aim is to prove that if \( X \) is a random symmetric matrix that is invariant under orthogonal transformations then such matrix \( X \) belongs to the GOE. We will do this using nothing more than the characteristic function.

1 Introduction

We will prove the characterization of the matrices that belong to the Gaussian Orthogonal Ensemble. This work was done as a project for the course of Random Matrices at the Centro de Investigación en Matemáticas, A.C. (CIMAT).

2 Notation

We will denote by \( M_{n \times m}(\mathbb{F}) \), the set of matrices of size \( n \times m \) over the field \( \mathbb{F} \). The matrix \( I_{d \times d} \) will denote the identity matrix of the set \( M_{d \times d}(\mathbb{F}) \). We will denote by \( O(d) \) the set of orthogonal matrices of \( M_{d \times d}(\mathbb{F}) \). This means, if \( O \in O(n) \) then \( O^T O = O O^T = I_{n \times n} \). A symmetric matrix \( X \) is a square matrix such that \( X^T = X \).

3 Background

We define the characteristic function of a random variable \( R \) to be a function \( \varphi_R : \mathbb{R} \to \mathbb{C} \) given by,

\[
\varphi_R(t) = \mathbb{E}[e^{itR}], \quad \forall t \in \mathbb{R}.
\]

In particular, given a random variable \( R \sim \mathcal{N}(\mu, \sigma^2) \), its characteristic function is given by:

\[
\varphi_R(t) = \mathbb{E}[e^{itR}] = e^{i \mu t - \frac{1}{2} \sigma^2 t^2}.
\]

Another important thing to remember is that if two random variables \( R_1, R_2 \) are such that their characteristic function coincide in every point, then their distributions are the same. In other words, if \( \varphi_X(t) = \varphi_Y(t) \), for every \( t \in \mathbb{R} \), then \( X \sim Y \).

Given a symmetric random matrix \( X \in \mathbb{S}_d \) its characteristic function is defined by:

\[
C_X : \mathbb{S}_d \to \mathbb{C}
\]

\[
C_X(M) = \mathbb{E}[\exp\{i \text{Tr}(X^T M)\}], \quad \forall M \in \mathbb{S}_d.
\]

Recall that the trace is invariant under cyclic permutations. From this, if \( A \in M_{d \times d}(\mathbb{R}) \) and for all \( M \in \mathbb{S}_d \),

\[
C_X(A^T MA) = \mathbb{E}[\exp\{i \text{Tr}(A^T X^T A M)\}] = \mathbb{E}[\exp\{i \text{Tr}(A^T X A M)\}] = \mathbb{E}[\exp\{i \text{Tr}([A^T X]^T M)\}] = C_{A^T X A}(M).
\]

* jose.sanchez@cimat.mx
† victor.amaya@cimat.mx
Notice that, if $A, B \in S_d$,

$$\text{Tr}(A^T B) = \sum_{j=1}^{d} (A^T B)_{jj} = \sum_{j,k=1}^{d} (A^T)_{jk} B_{kj} = \sum_{j,k=1}^{d} A_{kj} B_{kj}. \quad (1)$$

Furthermore, if $A, B \in S_d$,

$$\text{Tr}(A^T B) = \sum_{j=1}^{d} A_{jj} B_{jj} + 2 \sum_{1 \leq j < k \leq d} A_{jk} B_{jk}. \quad (2)$$

From this, note that

$$C_X(M) = E[\exp \{ i \text{Tr}(X^T M) \}] = E \left[ \exp \left\{ i \sum_{j=1}^{d} M_{jj} X_{jj} + 2i \sum_{1 \leq j < k \leq d} M_{kj} X_{kj} \right\} \right]. \quad (3)$$

Furthermore, if the entries of the matrix $X$ are independent, we get a concise formula of the characteristic function in terms of the characteristic functions of each entry:

$$C_X(M) = \prod_{j=1}^{d} E[\exp \{ i M_{jj} X_{jj} \}] \cdot \prod_{1 \leq j < k \leq d} E[\exp \{ 2i M_{kj} X_{kj} \}]. \quad (4)$$

It will be useful to keep this formula in mind when computing the value of the characteristic function on particular symmetric matrices.

### 4 The problem

Our aim is to prove the following theorem using nothing more than the characteristic function of a random matrix.

**Theorem 1.** Let $X$ be a random symmetric $d \times d$ matrix such that it is invariant under orthogonal transformations. Suppose that all the entries of $X$ are independent with finite variance. Then, there exist a matrix belonging to the Gaussian Orthogonal Ensemble $Y$, and real numbers $\mu \in \mathbb{R}$, $\sigma^2 \geq 0$ such that $X = \mu I_{d \times d} + \sigma^2 Y$.

We will divided our prove into three steps:

- Show that the elements inside the diagonal share the same distribution, and that all the elements outside the main diagonal share the same distribution.
- Show that is enough to prove the result for square $2 \times 2$ matrices.
- We will prove that the characteristic functions of the elements of the matrix $X$ correspond to normally distributed random variables.

**Proof:** Let’s proof every step.

**First step:** Let us define the matrix $A^t_{(k,j)}$ as the matrix that has the value $t$ in the positions $(k,j)$ and $(j,k)$ and zeros everywhere else, for $1 \leq k < j \leq d$ and $t \in \mathbb{R}$. If we calculate the the characteristic function of $A^t_{(k,j)}$ we obtain:

$$C_X(A^t_{(k,j)}(t)) = E \left[ \exp \left\{ i \text{Tr}((A^t_{(k,j)})^T X) \right\} \right] = E[\exp \{ 2it X_{kj} \}] = \phi_{X_{kj}}(2t). \quad (5)$$

Now, let $k \neq j$ and $l \neq g$. There exists a permutation matrix $P \in O(d)$ such that for all $t \in \mathbb{R}$, $P^t A^t_{(k,j)} P = A^t_{(l,g)}$. From the in-variance of $X$ with respect to orthogonal conjugations:
\( \varphi_{X_{tj}}(2t) = C_X(A_{[k,j]}^t) = C_X(P^T A_{[k,j]}^t P) = C_X(A_{[t,g]}^t) = \varphi_{X_{tg}}(2t), \quad \forall t \in \mathbb{R}. \)

It follows that \( X_{kj} \sim X_{tg} \). It means that all entries outside the main diagonal have the same distribution. In an analogous way, by conjugating we can prove that if \( 1 \leq j < k \leq d \), \( X_{kj} \sim -X_{jk} \). This implies the distribution of the entries outside of the diagonal are symmetric. This is because all entries of the matrix have finite second moment, \( \mathbb{E}(X_{kj}) = 0 \).

Now, let \( A_k^t = \text{diag}(t \cdot e_j) \), i.e. a matrix that has as only non-zero entry \( t \) in the position \((j,j)\), for \( 1 \leq j \leq d \) and \( t \in \mathbb{R} \). From this,

\[
C_X(A_k^t) = \mathbb{E} \left[ \exp \left\{ i \text{Tr}(A_k^t X) \right\} \right] = \mathbb{E} \left[ \exp \left\{ it X_{jj} \right\} \right] = \varphi_{X_{jj}}(t),
\]

for all \( 1 \leq j \leq d \) and \( t \in \mathbb{R} \). Now, given \( j \neq k \), there exists a permutation matrix \( Q \in O(d) \) such that \( Q^T A_j^t Q = A_k^t \). From this,

\[
\varphi_{X_{jj}}(t) = C_X(A_j^t) = C_X(Q^T A_j^t Q) = C_X(A_k^t) = \varphi_{X_{kk}}(t) \quad t \in \mathbb{R}.
\]

Thus \( X_{jj} \sim X_{kk} \) for all \( j \neq k \). Then, all entries in the diagonal have the same distribution.

**Second step:** In the last step, we have just proved that all entries in the diagonal have the same distribution, and the entries outside the diagonal also have the same distribution. From this, it is possible to find the distribution of all entries of the matrix \( X \) by finding the distribution of the entries of the sub-matrix,

\[
X_2 = \begin{pmatrix} X_{11} & X_{12} \\ X_{12} & X_{22} \end{pmatrix}.
\]

It can be proved that \( X_2 \) is invariant with respect to rotations in \( \mathbb{R}^2 \). For this, let \( \theta \in \mathbb{R} \) and

\[
Q^t_{\theta} = \begin{pmatrix} M_{\theta} & 0 \\ 0 & I_{(d-2) \times (d-2)} \end{pmatrix} \in \mathcal{M}_{d \times d}(\mathbb{R}),
\]

where,

\[
Q_{\theta} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix} \in \mathcal{M}_{2 \times 2}(\mathbb{R}).
\]

The matrix \( Q_{\theta}^t \) is an orthogonal matrix which rotates the first two coordinates and fixes the rest. Let \( M \in \mathbb{S}_2 \), where,

\[
M = \begin{pmatrix} a & b \\ b & d \end{pmatrix} \in \mathbb{S}_2.
\]

We can embed \( M \) into \( \mathbb{S}_d \) by filling all entries outside the \( 2 \times 2 \) first sub-matrix with zeros:

\[
M' = \begin{pmatrix} M & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{S}_d.
\]

We observe that,

\[
C_X(M') = \mathbb{E} \left[ \exp \left\{ i \text{Tr}(M'^T X) \right\} \right] = \mathbb{E} \left[ \exp \left\{ i(aX_{11} + 2bX_{12} + dX_{22}) \right\} \right] = C_X(M).
\]

By evaluating the characteristic function of \( X \) on \( M' \) and using the orthogonal in-variance, and considering that all entries outside the \( 2 \times 2 \) matrix in \( M' \) are zero,

\[
C_{X_2}(M) = C_X(M') = C_X((Q^t_{\theta} M') Q_{\theta}) = C_{X_2}(Q^t_{\theta} M Q_{\theta}).
\]
Third part: Since we have reduced the problem to the case of matrices of $2 \times 2$, we will keep $M$ and $Q_\theta$ to be the matrices defined in the previous step. We take $M_\theta = Q_\theta^T M Q_\theta$, a $2 \times 2$ symmetric matrix. Name the entries of $M_\theta$ as follow

$$M_\theta = \begin{pmatrix} A & B \\ B & D \end{pmatrix},$$

we can compute explicitly the values of $A, B, D$ in function the entries of $M$. We get the following:

$$A = \frac{a + d}{2} + \frac{a - d}{2} \cos(2\theta) - b \sin(2\theta),$$

$$B = \frac{a - d}{2} \sin(2\theta) + b \cos(2\theta),$$

$$D = \frac{a + d}{2} - \frac{a - d}{2} \cos(2\theta) + b \sin(2\theta).$$

These expressions are valid for every $\theta \in \mathbb{R}$. If we calculate the derivatives of the last expressions with respect to $\theta$, what we get is the following:

$$\frac{dA}{d\theta} = -2B, \quad \frac{dB}{d\theta} = A - D, \quad \frac{dD}{d\theta} = 2B$$

(1)

Now, since $X_2$ is invariant under rotations, we have that:

$$C_{X_2}(M_\theta) = C_{X_2}(M).$$

Computing the characteristic function on both matrices we obtain that,

$$\varphi_1(a)\varphi_2(2b)\varphi_1(d) = \varphi_1(A)\varphi_2(2B)\varphi_1(D),$$

(2)

where $\varphi_1$ represents the characteristic function of the random variable $X_{11}$ (which is the same of $X_{22}$ since the have the same distribution). And $\varphi_2$ is the characteristic function of $X_{12}$.

Since entries of $X_2$ have finite variance $\varphi_1$ and $\varphi_2$ are at least two times derivable at zero and $\phi_j'(0) = i\mathbb{E}(X_j)$ for $j = 1, 2$. We can derive equation (1) with respect to $\theta$, we get:

$$0 = \varphi_1'(A)\varphi_2(2B)\varphi_1(D)\frac{dA}{d\theta} + \varphi_1(A)\varphi_2'(2B)\varphi_1(D)\frac{dA}{d\theta} + \varphi_1(A)\varphi_2(2B)\varphi_1'(D)\frac{dD}{d\theta} + \varphi_1(A)\varphi_2(2B)\varphi_1(D)\frac{dD}{d\theta}$$

(3)

Note that since $\varphi_1(0) = \varphi_2(0) = 1$, and the are continuous at zero, there exist an open set around zero where these functions are not zero. So, we can rewrite equation (3) in the following way:

$$0 = \frac{\varphi_1'(A)}{\varphi_1(A)}\varphi_2(2B)\varphi_1(D)\frac{dA}{d\theta} + \frac{\varphi_2'(2B)}{\varphi_2(2B)}\varphi_1(A)\varphi_2(2B)\varphi_1(D)\frac{dA}{d\theta} + \frac{\varphi_1'(D)}{\varphi_1(D)}\varphi_1(A)\varphi_2(2B)\varphi_1(D)\frac{dD}{d\theta} + \frac{\varphi_1'(D)}{\varphi_1(D)}\varphi_1(A)\varphi_2(2B)\varphi_1(D)\frac{dD}{d\theta}.$$

Now, replacing the values we found in (1) and dividing by $-B(A - D)$ (for those values $B \neq 0, A \neq D$), we get that

$$1 \frac{\varphi_2'(2B)}{\varphi_2(2B)} = \frac{1}{A - D} \left[ \frac{\varphi_1'(A)}{\varphi_1(A)} - \frac{\varphi_1'(D)}{\varphi_1(D)} \right]$$

(4)

We observed that the left hand side of (4) only depends on $B$, whereas the right hand side depends in both $A$ and $D$. By this observation by can conclude that (4) is a constant. Then, there exists $k \in \mathbb{R}$ such that:

$$1 \frac{\varphi_2'(2B)}{\varphi_2(2B)} = -k,$$

Observe that, since $\varphi_2$ is the characteristic function of a symmetric random variable, the co-domain of $\varphi_2$ is contained in $\mathbb{R}$. This implies $k \in \mathbb{R}$. Now, lets take $x = 2B$. We have that

$$\varphi_2'(x) = -\frac{d}{dx} \varphi(x).$$

This is an easy real-valued ODE whose solution is given by $\varphi(x) = B_0 e^{-\frac{kx^2}{2}}$, $B_0$ are constants.
for some $B_0 \in \mathbb{R}$. Now, as $\varphi_2$ is a characteristic function it is true that $\varphi_2(0) = 1$. This give us as a result that $B_0 = 1$ so we have.

$$\varphi_2(x) = e^{-\frac{1}{2}k^2x^2}. $$

At first, we have this in a neighborhood of $0 \in \mathbb{R}$, but we can extend this solution to be for all $\mathbb{R}$. Observe that if $k < 0$ then $|\varphi_2(t)| \to \infty$ when $|t| \to \infty$. This leads to a contradiction since characteristic functions have bounded image. Then, if $k > 0$ and $\varphi_2$ is the characteristic function of a normal random variable with parameters $\mathcal{N}(0, k/2)$. Given $k = 0$, $\varphi_2$ is the characteristic function of a singular distribution at $0$.

On the other hand, taking $D = 0$ in equation (4), it follows that

$$\frac{1}{A} \left[ \frac{\varphi'_1(A)}{\varphi_1(A)} - \frac{\varphi'_1(0)}{\varphi_1(0)} \right].$$

We know that $\varphi_1(0) = 1$ and that $\varphi'(0) = i\mu$, where $\mu = \mathbb{E}(X_{11})$. So,

$$\frac{\varphi'_1(A)}{\varphi_1(A)} = -Ak + i\mu \quad \Rightarrow \quad \varphi'_1(A) = (-Ak + i\mu)\varphi_1(A).$$

The solution to this ODE is $\varphi_1(x) = e^{i\mu x - \frac{k^2}{4}x^2}$, which is nothing more than the characteristic function of a random variable with distribution $\mathcal{N}(\mu, k)$.

Since $k$ is non-negative we can write it as $k = 2\sigma^2$. So we can write that characteristic functions in the following way:

$$\varphi_1(x) = e^{i\mu x - \sigma^2x^2}, \quad \varphi_2(x) = e^{-\sigma^2\frac{x^2}{2}}.$$

which means that $X_{11} \sim \mathcal{N}(\mu, 2\sigma^2)$ and $X_{12} \sim \mathcal{N}(0, \sigma^2)$.

With this, we proved that $X_{jj} \sim \mathcal{N}(\mu, 2\sigma^2)$ and $X_{kj} \sim \mathcal{N}(0, \sigma^2)$. Now, if the take a GOE matrix $Y$, i.e., $Y_{ii} \sim \mathcal{N}(0, 2)$ and $Y_{jk} \sim \mathcal{N}(0, 1)$, for every $j \neq k$, we conclude that:

$$X \sim \sigma^2Y + \mu I_d.$$

We see that $X$ is clearly invariant under orthogonal conjugations. □

5 Conclusions

There are important remarks that are worth stressing about this proof. First, we find interesting that it unfolded by using just elementary knowledge of basic properties of characteristic functions, differential equations and linear algebra. It is a basic proof for this result.

Second, it states a generalization of the standard result. This result is proved usually for matrices that have zero-mean entries. This proof allow us to study general matrices that are invariant with respect to orthogonal conjugations.

The independence of the entries is key for this proof. One example of a distribution that is invariant under orthogonal conjugation is the Haar distribution on $O(d)$. This distribution does not hold the entry-wise independence since orthogonal matrices satisfy relations between them. In terms of this work, it is key to write the characteristic function of the matrix in terms of the product of the characteristic function of each entry.

One possible extension of this project is to find the distribution of Wishart matrices by applying the same characteristic-function approach.

References

[1] Yi-Kai Liu. Mathematics Junior Seminar, Spring 2001. Princeton University. 
http://web.math.princeton.edu/mathlab/projects/ramatrices/yl/randmtx.PDF.

[2] Measure Theory and Probability Theory. Krishna B. Athreya and Soumendra N. Lahiri. Springer Text in Statistics.