Notas del curso de Matrices Aleatorias

Impartido por:

Dr. Mario Alberto Díaz Torres

Dr. Víctor Manuel Pérez-Abreu Carrión

Dr. Carlos Vargas Obieta

Apoyadas en las notas de Daniel Perales Anaya del curso del 2017.

Elaboradas por Saúl Rogelio Mendoza Jacobo.

2018

Contenido

Capítulo 1

Matrices Aleatorias

1.1. Preliminares de Álgebra Lineal

Antes de definir y estudiar el concepto principal del curso establezcamos algo de notación y algunos conceptos que serán recurrentes en lo siguiente.

Denotamos por $\mathbb{M}_{d\times e}(\mathbb{F})$ el espacio lineal de matrices $d \times e$ con entradas en el campo $\mathbb{F}$; si $d = e$, simplemente escribimos $\mathbb{M}_d(\mathbb{F})$. En este curso se considerarán sólo los casos $\mathbb{F} = \mathbb{R}$ y $\mathbb{F} = \mathbb{C}$. Para $B \in \mathbb{M}_{d\times e}(\mathbb{F})$, B^* denota la *matriz adjunta* de B, en el caso real esto es $B^* = B^T$ y en el caso complejo $B^* = (\bar{b}_{ji})$, si $B = (b_{ij})$. Decimos que una matriz cuadrada $A \in \mathbb{M}_d(\mathbb{F})$ es *semidefinida positiva* si $\forall y \in \mathbb{F}^d,\ y^*Ay \geq 0$. Aprovechando esta última definición, si A, B matrices cuadradas, decimos que $A \geq B$ si $A - B$ es semidefinida positiva.

Para $A, B \in \mathbb{M}_{d\times e}(\mathbb{F})$ arbitrarios, B^*A es matriz cuadrada de $e \times e$ y por tanto podemos definir el *producto interior de Frobenius*:

$$\langle A, B\rangle = Tr(B^*A) = \sum_{i=1}^{d}\sum_{j=1}^{e} \bar{b}_{ij}a_{ij},$$

lo cual induce la norma

$$\|A\| = \langle A, A\rangle^{1/2} = \left[\sum_{i=1}^{d}\sum_{j=1}^{e}|a_{ij}|^2\right]^{1/2}.$$

Decimos que $\mathcal{Q} \subset \mathbb{M}_{d\times e}(\mathbb{F})$ es *un cono* si es cerrado bajo suma y multiplicación por reales positivos, i.e., si para todo $A, B \in \mathcal{Q}$ y $r \in \mathbb{R}^+$, se cumple que $A + B \in \mathcal{Q}$ y $rA \in \mathcal{Q}$.

Observación 1.1. Si tomamos un subespacio $\mathbb{W}$ de $\mathbb{M}_{d\times e}(\mathbb{F})$ de dimensión $n \leq d\times e$, ese subespacio es claramente isomorfo a $\mathbb{F}^n$; esto, junto con que $\mathbb{C}^n \cong \mathbb{R}^{2n}$ hace que identifiquemos cualquier subespacio de $\mathbb{M}_{d\times e}(\mathbb{F})$ (en ambos casos $\mathbb{C}$ y $\mathbb{R}$) con $\mathbb{R}^n$ para algún n, es decir, si $\mathbb{W}$ subespacio de $\mathbb{M}_{d\times e}(\mathbb{C})$ de dimensión n, entonces $\mathbb{W} \cong \mathbb{R}^{2n}$ y si en cambio $\mathbb{W}$ subespacio de $\mathbb{M}_{d\times e}(\mathbb{R})$ de dimensión n, entonces $\mathbb{W} \cong \mathbb{R}^n$.

Como primer ejemplo identificamos $\mathbb{W} = \mathbb{M}_{d\times e}(\mathbb{C})$ con $\mathbb{R}^{2(d\times e)}$ y $\mathbb{W} = \mathbb{M}_{d\times e}(\mathbb{R})$ con $\mathbb{R}^{d\times e}$.

Además, como todo espacio normado de dimensión finita es completo, podemos decir que $(\mathbb{M}_{d\times e}(\mathbb{F}), \|\cdot\|)$ es un espacio de Banach, ya sea que consideremos $\|\cdot\|$ la norma de Frobenius o cualquier norma de $\mathbb{R}^n$, con $n = d \times e$ en el caso real y $n = 2(d \times e)$ en el caso complejo; esto último porque en espacios normados de dimensión finita todas las normas son equivalentes.

Definimos ahora el producto tensorial de matrices.
Si $C = C_{d\times e}$ y $D = D_{m\times n}$ definimos su producto tensiorial o *producto de Kronecker* a la matriz a bloques $C \otimes D$ de dimensión $dm \times en$ definida por:

$$X = \begin{pmatrix} C_{11}D & C_{12}D & \cdots & C_{1e}D \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ C_{d1}D & C_{d2}D & \cdots & C_{de}D \end{pmatrix},$$

Propiedades del producto tensorial de matrices
En el caso de matrices cuadradas $d = e$ y $m = n$:

- $Tr(C \otimes D) = Tr(C)Tr(D)$.
- $\det(C \otimes D) = \det(C)^d \det(D)^n$.

Donde la traza y el determinante están definidos en sus respectivos espacios.

1.2. Primeras Definiciones

Denotemos por $\mathcal{B}(\mathbb{M}_{d\times e})$ los conjuntos de Borel de $\mathbb{M}_{d\times e}(\mathbb{C})$ o $\mathbb{M}_{d\times e}(\mathbb{R})$. Definimos ahora el concepto principal de este curso.

Definición 1.1. Sea $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y $\mathcal{Q} \subset \mathbb{M}_{d\times e}(\mathbb{F}), \mathcal{Q} \in \mathcal{B}(\mathbb{M}_{d\times e})$. Una función $X : \Omega \to \mathcal{Q}$ es una matriz aleatoria si

$$X^{-1}(A) \in \mathcal{F} \qquad \forall A \in \mathcal{B}(\mathcal{Q}) = \mathcal{Q} \cap \mathcal{B}(\mathbb{M}_{d\times e}).$$

O bien, X es una función $\mathcal{B}(\mathbb{M}_{d\times e}) \setminus \mathcal{F}$-medible.

Por un resultado de Teoría de la Medida, si $X = (X_{ij}) \in \mathcal{Q}$, entonces X es matriz aleatoria con valores en $\mathcal{Q}$ si y sólo si X_{ij} es una variable aleatoria (compleja o real) para todo $i = 1, \ldots, d$, $j = 1, \ldots, e$.

Observación 1.2. Los $\mathcal{Q}$ de interés en matrices aleatorias y su identificación con $\mathbb{R}^m$ para algún m son:

1. $\mathcal{Q} = \mathbb{M}_{d\times e}(\mathbb{C}) \cong \mathbb{R}^{2de}$, $m = 2de$.
2. $\mathcal{Q} = \mathbb{M}_{d\times e}(\mathbb{R}) \cong \mathbb{R}^{de}$, $m = de$.

   Si $d = e$**:**
3. $\mathcal{Q} = \mathbb{H}_d(\mathbb{C}) \cong \mathbb{R}^{d^2}$ matrices hermitianas ($A \in \mathbb{H}_d(\mathbb{C})$ sii $A^* = A$). $\mathbb{H}_d$ es subespacio lineal de $\mathbb{M}_d(\mathbb{C})$.
4. $\mathcal{Q} = \mathbb{S}_d(\mathbb{R}) \cong \mathbb{R}^{\frac{d(d+1)}{2}}$ matrices simétricas ($A \in \mathbb{S}_d(\mathbb{C})$ sii $A^T = A$). $\mathbb{S}_d$ es subespacio lineal de $\mathbb{M}_d(\mathbb{R})$.

5. $\mathcal{Q} = \mathbb{H}_d^+ \subset \mathbb{R}^{d^2}$ matrices semidefinidas positivas (con entradas complejas). $\mathcal{Q}$ es un cono.
6. $\mathcal{Q} = \mathbb{S}_d^+ \subset \mathbb{R}^{d^2}$ matrices semidefinidas positivas (con entradas reales), también es un cono.
7. $\mathcal{Q} = \mathbb{U}(d)$ el grupo unitario [$\mathbb{U}(d) = \{U \in \mathbb{M}_d(\mathbb{C}) : U^*U = UU^* = I_d\}$]
8. $\mathcal{Q} = \mathbb{O}(d)$ el grupo ortogonal [$\mathbb{O}(d) = \{O \in \mathbb{M}_d(\mathbb{R}) : O^TO = OO^T = I_d\}$]

Definición 1.2. La distribución en $\mathcal{Q}$ de la matriz aleatoria $X \in \mathcal{Q}$ es la medida de probabilidad

$$\mu_X(A) = \mathbb{P}(X \in A), \qquad A \in \mathcal{B}(\mathcal{Q}).$$

La función característica de X en $\mathcal{Q}$ es

$$c_X(T) = \mathbb{E}e^{iRe(Tr(T^*X))} \qquad \forall T \in \mathcal{Q},$$

donde Tr es el operador traza (se denota con mayúscula, pues tr se usará más adelante para denotar el operador $\frac{1}{n}Tr$ que llamaremos *traza normalizada*).

Ejemplo 1.1. Sea X variable aleatoria compleja, $X = Re(X) + iIm(X) =: X_1 + iX_2$, y X_1, X_2 son v.a.'s reales. $T = Re(T) + iIm(T) =: t_1 + it_2$, con $t_1, t_2 \in \mathbb{R}$. Son elementos de $\mathbb{M}_{1\times 1}(\mathbb{C})$ así que aquí la traza es la identidad. Entonces:

$$Re(\langle X, T\rangle) = Re(Tr(T^*X)) = Re(T^*X) = Re((t_1 - it_2)(X_1 + iX_2)) = t_1X_1 + t_2X_2,$$

que es el producto punto de (t_1, t_2) con (X_1, X_2), las representaciones de T y X en $\mathbb{R}^2$; es decir $c_X(T)$ es en este caso la función característica usual del vector (X_1, X_2), evaluado en el vector determinista (t_1, t_2). Esto explica porqué tomar Re en la definición de función característica.

Definición 1.3. Decimos que la matriz aleatoria X en $\mathcal{Q}$ tiene densidad si existe $f : \mathcal{Q} \to [0, \infty)$ tal que

$$\mathbb{P}(X \in A) = \int_{\mathcal{Q}\cap A} f(x)dx \qquad \forall A \in \mathcal{B}(\mathcal{Q})$$

donde $dx = \prod_{j=1}^m dx_j$ con $\mathcal{Q} \cong \mathbb{R}^m$ y

$$\int_{\mathcal{Q}} f(x)dx = 1.$$

Observación 1.3. 1. Matrices aleatorias en $\mathbb{U}(d)$ y $\mathbb{O}(d)$ no pueden tener densidad. Ya que ambos tienen volumen cero en los respectivos espacios $\mathbb{M}_d(\mathbb{C})$ y $\mathbb{M}_d(\mathbb{R})$.

2. Cuando se trabaja con $\mathbb{H}_d^+$ y $\mathbb{S}_d^+$ también se usa la transformada de Laplace,

$$\mathcal{L}(T) = \mathbb{E}e^{-Tr(T^*X)} \qquad \forall T \in \mathbb{H}_d^+ \text{ o } \forall T \in \mathbb{S}_d^+.$$

Ejemplo 1.2. 1) Sean $Z_{ij} \sim N(0,1)$ v.a. independientes para $i = 1, \dots, d$, $j = 1, \dots, e$. Sea $Z = (Z_{ij}) \in M_{d\times e}(\mathbb{R})$, llamamos a esta Ginibre Normal rectangular.

1') Como segundo ejemplo tenemos $Z = (Z_{ij}) \in M_{d\times e}(\mathbb{C})$, con Z_{ij} variables aleatorias independientes normales complejas, i.e., $Re(Z_{ij}) \sim N(0, 1/2)$ y $Im(Z_{ij}) \sim N(0, 1/2)$, con $Re(Z_{ij})$ independiente de $Im(Z_{ij})$. $Z_{ij} = Re(Z_{ij}) + iIm(Z_{ij})$.

Breviario Cultural

a) El caso 1) pero con $d = e = 2$ fue estudiado por Fisher en los 20's.

$$Z = \begin{pmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{pmatrix},$$

y más generalmente, $d = 2$ y $e = n$,

$$Z = \begin{pmatrix} Z_{11} & Z_{12} & \cdots & Z_{1n} \\ Z_{21} & Z_{22} & \cdots & Z_{2n} \end{pmatrix},$$

que es una muestra aleatoria de un vector de dos características. Wishart (1928) consideró el caso $d = p$, muestra de vector con p características.

b) El caso 1) lo consideró Von Neumann (1947) en álgebra numérica. Definió el *número de condición de una matriz* Z como

$$\kappa_2(Z) = \sqrt{\frac{\lambda_{max}(Z^*Z)}{\lambda_{min}(Z^*Z)}}, \tag{1.1}$$

con $\lambda_{max}(Z^*Z)$ el eigenvalor máximo de Z^*Z, y $\lambda_{min}(Z^*Z)$ el mínimo. Lo anterior está bien definido tanto para el ejemplo 1) como el 1') pues en esos casos Z^*Z es no singular con probabilidad 1, como lo veremos más adelante.

En la tesis doctoral de Edelman (1989, [Ede89]), podemos ver para el caso 1), el siguiente resultado:

$$\mathbb{P}\left[\frac{\kappa_2(Z)}{g(d,e)} > x\right] \leq \begin{cases} e^{-1/x^2}, \\ \frac{2}{\pi}\left(\frac{c}{x}\right)^{h(d,e)} \end{cases}. \tag{1.2}$$

c) El ejemplo 1') es el modelo básico en comunicación inalámbrica con d antenas transmisoras y e antenas receptoras y Z_{ij} es el coeficiente de propagación entre la antena transmisora i y la receptora j.

Proposición 1.1. Sean $Z_{ij} \sim N(0,1)$ v.a. independientes para $i = 1, \ldots, d$, $j = 1, \ldots, e$. Sea $Z = (Z_{ij}) \in M_{d\times e}(\mathbb{R})$. Entonces

1) $c_Z(T) = \mathbb{E}e^{iTr(T^tZ)} = e^{-\frac{1}{2}\|T\|^2} = e^{-\frac{1}{2}Tr(T^tT)}$.

2) Z tiene densidad

$$f_Z(z) = \left(\frac{1}{\sqrt{2\pi}}\right)^{de} e^{-\frac{1}{2}\|z\|^2} \qquad \forall z \in \mathbb{M}_{d\times e}(\mathbb{R}) \cong \mathbb{R}^{d\times e}.$$

3) Si $O \in \mathbb{O}(d)$ no aleatoria, $OZ \stackrel{d}{=} Z$ y $\forall P \in \mathbb{O}(e)$, $ZP \stackrel{d}{=} Z$.

4) En general, $\mathbb{P}(Rango(Z) = \min(d, e)) = 1$.

Demostración.

1) Sea $T = (T_{ij})$, veamos que

$$\mathbb{E}e^{iTr(T^tZ)} = \mathbb{E}e^{i\left(\sum\limits_{i=1}^{d}\sum\limits_{j=1}^{e} Z_{ij}T_{ij}\right)} = \mathbb{E}\prod_{i=1}^{d}\prod_{j=1}^{e} e^{iT_{ij}Z_{ij}} \overset{\text{(Indep.)}}{=} \prod_{i=1}^{d}\prod_{j=1}^{e} \mathbb{E}e^{iT_{ij}Z_{ij}}.$$

Observación: hasta aquí no hemos usado la distribución de Z_{ij}.

$$\overset{\text{(N(0,1))}}{=} \prod_{i=1}^{d}\prod_{j=1}^{e} e^{-\frac{1}{2}T_{ij}^2} = e^{-\frac{1}{2}Tr(T^tT)} = e^{-\frac{1}{2}\|T\|^2}.$$

2) Como las $Z_{ij} \sim N(0, 1)$ y además todas las entradas de la matríz son independientes entonces el vector

$$Vec(Z) = (Z_{11}, Z_{12}, \ldots, Z_{1e}, Z_{21}, \ldots, Z_{2e}, \ldots, Z_{d1}, \ldots Z_{de}) \in \mathbb{R}^{de},$$

se distribuye como una normal multivariada (de dimensión de) con media 0 y matriz de covarianzas I_{de}, y por lo tanto tiene densidad

$$f(z_{11}, \ldots, z_{1e}, \ldots, z_{d1}, \ldots, z_{de}) = \left(\frac{1}{\sqrt{2\pi}}\right)^{de} e^{-\frac{1}{2}\left(\sum_{i=1}^{d}\sum_{j=1}^{e} z_{ij}^2\right)} = \left(\frac{1}{\sqrt{2\pi}}\right)^{de} e^{-\frac{1}{2}\|z\|^2}.$$

Por el isomorfismo $\mathbb{M}_{d\times e}(\mathbb{R}) \cong \mathbb{R}^{de}$, concluimos que

$$f_Z(z) = \left(\frac{1}{\sqrt{2\pi}}\right)^{de} e^{-\frac{1}{2}\|z\|^2}, \qquad \forall z \in \mathbb{M}_{d\times e}(\mathbb{R}).$$

3) Obtengamos la función característica de UZ.

$$\mathbb{E}e^{iTr(T^t(UZ))} = \mathbb{E}e^{iTr((U^tT)^tZ)}.$$

Usando el inciso 1), tomando T como $U^tT \in \mathbb{M}_{d\times e}(\mathbb{R})$, entonces

$$= e^{-\frac{1}{2}Tr((U^tT)^t(U^tT))} = e^{-\frac{1}{2}Tr(T^tUU^tT)} = e^{-\frac{1}{2}Tr(T^tT)} = \mathbb{E}e^{iTr(T^tZ)}, \qquad \forall T \in \mathbb{M}_{d\times e}(\mathbb{R}),$$

que es la función característica de Z y por la unicidad de la función característica, tenemos $UZ \overset{d}{=} Z$. El otro caso es análogo.

4) Se obtiene en la sección 1.3.

■

El resultado anterior puede extenderse al caso $Z_{ij} \sim N(0, \sigma^2)$. Para el caso complejo tenemos un resultado similar.

Proposición 1.2. Sea Z_{ij} con $Re(Z_{ij}) \sim Im(Z_{ij}) \sim N(0, \frac{1}{2})$ v.a. independientes para $i = 1, \ldots, d$, $j = 1, \ldots, e$. Sea $Z = (Z_{ij}) \in M_{d\times e}(\mathbb{C})$. Entonces

1) $c_Z(T) = \mathbb{E}e^{iRe[Tr(T^*Z)]} = e^{-\frac{1}{4}Tr(T^*T)} = e^{-\frac{1}{4}\|T\|^2}, \ \forall \in \mathbb{M}_{d\times e}(\mathbb{C})$.

2) Z tiene densidad
$$f_Z(z) = \frac{1}{\pi^{de}} e^{-Tr(z^*z)}, \qquad \forall z \in \mathbb{M}_{d\times e}(\mathbb{C}).$$

3) Si $U \in \mathbb{U}(d)$ no aleatoria, $UZ \stackrel{d}{=} Z$ y $\forall\ V \in \mathbb{U}(e)$, $ZV \stackrel{d}{=} Z$.

4) En general, $\mathbb{P}(Z \text{ tiene rango completo}) = 1$.

Demostración.

1) Cálculo muy similar al de la proposición anterior.

2) Se obtiene de manera similar a la de la proposición anterior con la identificación $\mathbb{M}_{d\times e}(\mathbb{C}) \cong \mathbb{C}^{de} \cong \mathbb{R}^{2de}$.

3) Queremos demostrar que $\forall\ U \in \mathbb{U}(d)$ no aleatoria, $UZ \stackrel{d}{=} Z$.
$$\mathbb{E}e^{iReTr(T^*(UZ))} = \mathbb{E}e^{iReTr((U^*T)^*Z)}.$$
Usando el primer inciso, tomando T como $U^*T \in \mathbb{M}_{d\times e}(\mathbb{C})$, entonces
$$= e^{-\frac{1}{4}Tr((U^*T)^*(U^*T))} = e^{-\frac{1}{4}Tr(T^*UU^*T)} = e^{-\frac{1}{4}Tr(T^*T)} = \mathbb{E}e^{iReTr(T^*Z)}, \qquad \forall T \in \mathbb{M}_{d\times e}(\mathbb{C}).$$
Como la función característica determina a la distribución, tenemos $UZ \stackrel{d}{=} Z$. El otro caso es análogo.

4) Se obtiene en la siguiente sección.

∎

1.3. Singularidad de matrices aleatorias

Comenzamos con dos definiciones que serán recurrentes a lo largo de estas notas.

Definición 1.4. 1. Si X es matriz aleatoria en $M_{d\times d}(\mathbb{C})$ con entradas independientes, la llamamos *Matriz de Ginibre* (en caso no cuadrado lo llamamos *Matriz de Ginibre rectangular*).

2. Se dice que $X = (X_{ij}) \in M_{d\times d}(\mathbb{F})$ es *Matriz de Wigner* si

   i) (Caso real) X es real simétrica y $\{X_{ij} : 1 \leq i \leq j \leq d\}$ son variables aleatorias independientes.

   ii) (Caso complejo) X es hermitiana compleja y $\{X_{ij} : 1 \leq i \leq j \leq d\}$ son variables aleatorias independientes.

Nota 1.1. A veces se pide para una matriz de Wigner que tenga segundo momento finito y que encima de la diagonal tenga un medio de la varianza de la diagonal. También es común pedir segundo momento y simetría de las distribuciones.

Lema 1.1 (No singularidad de matrices aleatorias con densidad)**.** Sea X matriz aleatoria que toma valores en $\mathcal{Q} \cong \mathbb{R}^m$, con función de densidad $f_X(x)$ con respecto a la medida de Lebesgue en $\mathbb{R}^m$. Entonces,

$$\mathbb{P}(rango(X) \text{ completo}) = 1.$$

Observación 1.4. El conjunto

$$R(X) = \{X \text{ tiene rango completo}\} = \{\omega \in \Omega : X(\omega) \text{ tiene rango completo}\} \in \mathcal{F},$$

es decir, es medible.

En efecto, sea X de dimensión $p \times q$ con $q \leq p$. Entonces

$$R(X) = \{\det(X^*X) \neq 0\} = \{\det(X^*X) \in \mathbb{R} \backslash \{0\}\},$$

y $\mathbb{R} \backslash \{0\} \in \mathcal{B}(\mathbb{R})$, y como para una matriz cuadrada, $\det(A)$ es un polinomio en las entradas (y por tanto una función continua), se sigue que $R(X) \in \mathcal{F}$ y $(R(X))^C \in \mathcal{F}$.

Demostración.

Como X tiene densidad f_X,

$$\mathbb{P}(X \text{ tiene rango completo}) = \int \cdots \int_{\det(x^*x)=0} f_X(x) dx.$$

El resultado se sigue del siguiente resultado de análisis matemático 2: Sea $P(x_1, \ldots, x_m)$ un polinomio de m variables. Entonces $\{(x_1, \ldots, x_m) \in \mathbb{R}^m : P(x_1, \ldots, x_m) = 0\}$ tiene medida de Lebesgue 0.

■

Ejemplo 1.3. Si

$$X = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix},$$

donde X_{ij} son v.a.i. con distribución discreta Bernoulli $p = \frac{1}{2}$, entonces tenemos que $\det(X) = X_{11}X_{22} - X_{12}X_{21}$ y que

$$\mathbb{P}(det(X) = 0) = \mathbb{P}(X_{11}X_{22} - X_{12}X_{21} = 0) > 0.$$

El caso en que X_{11} tiene distribución continua puede resolverse dado el hecho de que si C y D son v.a.i. y C tiene distribución continua, entonces $C + D$ tiene distribución continua y se discute en la subsección.

Definición 1.5. Una matriz aleatoria $X = (X_{ij}) \in \mathbb{M}_d(\mathbb{R})$ es *del tipo* M si $\{X_{11}, \ldots, X_{dd}\}$ y $\{X_{ij}, i \neq j\}$ son independientes y $X_{11}, \ldots, X_{dd}$ son v.a.i. y tienen distribución continua.

Teorema 1.1 (Teorema 5, Manrique(2017))**.** Si X es matriz aleatoria de tipo M entonces

$$\mathbb{P}(\det(X) \neq 0) = 1.$$

Ejercicio 1.1. Demostrar el teorema anterior para X con entradas en los complejos.

Solución: Procedemos por inducción, para $d = 1$ tenemos que $\mathbb{P}(\det(X) \neq 0) = \mathbb{P}(X \neq 0) = 1$, porque X es continua. Ahora, si suponemos que el teorema es cierto para $k \leq d-1$, lo demostraremos para d. Observemos que

$$\det(X^{(d)}) = \xi_{1,1} \det(X^{(d-1)}) + \sum_{j=2}^{d} \xi_{1,j} c_{1,j},$$

dónde $X^{(d)}$ es una matriz de dimensión d que cumple las hipótesis y $X^{(d-1)}$ es la submatriz que se obtiene de quitar la primera fila y columna, y $c_{1,j}$ es el $(1,j)$ cofactor de $X^{(d)}$. Por hipótesis de inducción tenemos que $\det(X^{(d-1)}) \neq 0$. Entonces tendremos:

$$\begin{aligned}\mathbb{P}(\det(X^{(d)}) \neq 0) &= \mathbb{P}\left(\xi_{1,1} \det(X^{(d-1)}) + \sum_{j=2}^{d} \xi_{1,j} c_{1,j} \neq 0\right) = \mathbb{P}\left(\xi_{1,1} \neq -\frac{\sum_{j=2}^{d} \xi_{1,j} c_{1,j}}{\det(X^{(d-1)})}\right) \\ &= \mathbb{E}\left(\mathbb{P}\left(\xi_{1,1} \neq -\frac{\sum_{j=2}^{d} \xi_{1,j} c_{1,j}}{\det(X^{(d-1)})}\right) |\xi_{i,j} \text{ con}(i,j) = [d]^2 \backslash (1,1)\right) = \mathbb{E}(1) = 1.\end{aligned}$$

Corolario 1.1.

1. Si $X = (X_{ij})$ donde X_{ij} v.a.i. con distribución continua, entonces X es de tipo M y $\mathbb{P}(\det(X) \neq 0) = 1$.
2. Una matriz de Ginibre con elementos en la diagonal con distribución continua es no singular con probabilidad 1.
3. Una matriz de Wigner con elementos en la diagonal con distribución continua es no singular con probabilidad 1.
4. Se tiene que se cumple el último inciso en las Proposiciones 1.1 y 1.2.

Observación 1.5. Si $\tilde{X}$ matriz aleatoria de Ginibre con un elemento en cada renglón y en cada columna con distribución continua entonces

$$\mathbb{P}[\det(\tilde{X}) \neq 0] = 1.$$

En efecto, sea P una matriz de permutación tal que $Y = P\tilde{X}$ tiene en la diagonal las variables aleatorias con distribución continua. Entonces

$$\det(Y) = \det(P) \det(\tilde{X}),$$

pero como Y es de tipo M y $\det(P) = 1$ por ser ortogonal; concluimos que

$$0 \neq \det(Y) = \det(P) \det(\tilde{X}) = \det(\tilde{X}),$$

con probabilidad 1.

Nos planteamos ahora la pregunta: ¿Qué ocurre cuando tenemos sólo k variables continuas en la diagonal? ¿tendremos que la matriz tiene rango k?. Tal cuestión se aborda en la siguiente subsección.

1.3.1. Sobre el Rango de Matrices Aleatorias

A continuación se exploran algunas relaciones entre el número de entradas con distribución continua en una matriz aleatoria y el rango de ésta. Se demuestra una generalización del teorema 1.1 para matrices rectangulares con sólo $k \leq \min(d, e)$ entradas continuas en la diagonal principal.

Definición 1.6. Sea $X = (X_{ij}) \in \mathbb{M}_{d\times e}(\mathbb{R})$ matriz aleatoria y $k \leq \min(d, e)$. Decimos que X es *del tipo* M_k si tiene k variables aleatorias independientes continuas en su diagonal principal, digamos $X_{i_1 i_1}, \ldots, X_{i_k i_k}$ y además $\{X_{i_1 i_1}, \ldots, X_{i_k i_k}\}$ y $\{X_{ij} : (i, j) \notin \{(i_1, i_1), \ldots, (i_k, i_k)\}\}$ son conjuntos independientes.

Usaremos la siguiente caracterización del rango de una matriz.

Lema 1.2. El rango de una matriz $A \in M_{d\times e}(\mathbb{F})$ es el natural más grande r tal que A tiene una submatriz B de $r \times r$ con determinante no cero.

Observación 1.6. Otra manera de enunciar lo anterior es:
El rango de una matriz $M \in \mathbb{M}_{d\times e}(\mathbb{R})$ es el natural r más grande tal que M tiene un menor de orden r no cero.

Teorema 1.2. Sea $A \in M_{d\times e}(\mathbb{R})$ matriz aleatoria de tipo M_k. Entonces $rank(A) \geq k$ casi seguramente, es decir $\mathbb{P}(rank(A) \geq k) = 1$.

Demostración. Aplicando operaciones elementales de cambio de filas y/o columnas a A podemos llevarla a una matriz de la forma

$$A' = \begin{pmatrix} B & C_1 \\ C_2 & C_3 \end{pmatrix},$$

donde B es la submatriz $k \times k$, que tiene en la diagonal a las variables aleatorias continuas independientes y por el lema de agrupamientos (subclases de clases independientes son independientes) tenemos que B es del tipo M.

Podemos usar entonces el teorema 1.1 para afirmar que $\det(B) \neq 0$ con probabilidad 1 y por tanto, por el lema 1.2 concluimos que $rank(A') \geq k$ con probabilidad 1; esto último ya que B es una submatriz con determinante no cero c.s. y el rango (lema 1.2) es el máximo valor m tal que existe una submatriz no singular de $m \times m$. Como las operaciones elementales de cambio de fila y/o columna no cambian el rango tenemos finalmente que $rank(A) = rank(A')$ y por tanto:

$$\mathbb{P}(rank(A) \geq k) = \mathbb{P}(rank(A') \geq k) = 1.$$

∎

Observación 1.7. Otra manera de probar lo anterior es si damos por hecho la observación 1.6 y llevamos a A mediante operaciones elementales de cambio de fila a la forma:

$$A = \begin{bmatrix} A' & B_1 \\ B_2 & B_3 \end{bmatrix} \tag{1.3}$$

con A' una matriz $k \times k$ del tipo $\mathcal{M}$. Así, por el teorema 1.1 se tiene que $\mathbb{P}(\det(A') \neq 0) = 1$.

Por otra parte, si r es el número de menores de orden k de A y éstos son $l_1, \ldots, l_r$, se tiene que, por la observación 1.6,

$$\mathbb{P}(rank(A) \geq k) = \mathbb{P}(\cup_i \{l_i \neq 0\}).$$

Finalmente, dado que

$$\mathbb{P}(\cup_i \{l_i \neq 0\}) \geq \mathbb{P}(\det(A') \neq 0),$$

se concluye lo deseado. i.e., que $\mathbb{P}(rank(X) \geq k) = 1$.

Corolario 1.2. Si A matriz de Ginibre con k variables continuas como entradas pero al menos una variable continua por cada fila y cada columna de una colección de k filas y columnas, cumple que $\mathbb{P}(rank(A) \geq k) = 1$.

Demostración. Se sigue de tomar una matriz de permutación como en la observación 1.5. ∎

Corolario 1.3 (teorema 1.1 para matrices rectangulares)**.** Si la matriz aleatoria $A \in M_{d\times e}(\mathbb{R})$ es del tipo $M_{\min(e,d)}$. Entonces A es de rango completo con probabilidad 1.

Por último abordamos la pregunta: **¿podemos decir más?¿por ejemplo que $\mathbb{P}(rank(A) = k) = 1$ con tales hipótesis?**

Para responder a tal pregunta consideremos la matriz aleatoria

$$X = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix},$$

con X_{12}, X_{21}, X_{22} i.i.d. con distribución Rademacher y X_{11} con distribución continua e independiente de las demás entradas. Entonces:

$$\mathbb{P}(det(X) \neq 0) = \mathbb{P}(X_{11}X_{22} - X_{12}X_{21} \neq 0) = \mathbb{P}(X_{11} \neq \frac{X_{12}X_{21}}{X_{22}}) = 1.$$

Es decir, aunque X tenga $k = 1$ variables continuas en la diagonal

$$\mathbb{P}(rank(X) = 2 > 1 = k) = 1.$$

Así que la respuesta es **no**, no podemos decir más que $\mathbb{P}(rank(A) \geq k) = 1$ con las hipótesis del teorema.

1.3.2. Más sobre Singularidad

En esta subsección se consideran otros aspectos sobre singularidad de matrices aleatorias (clase de Paulo Manrique 2017). Consideremos $A_{n\times n} = (y_{ij})_{1\leq i,j\leq n}$ donde y_{ij} son v.a. El problema de interés es encontrar $\mathbb{P}(det(A_{n\times n}) = 0)$, esto se puede responder si conocemos la distribución de $det(A_{n\times n})$, $|det(A_{n\times n})|$ o incluso si conocemos asintóticamente cosas del tipo

$$a_n = \frac{\frac{1}{n}\log|\det(A_{n\times n})| + b_n}{c_n}.$$

Para el caso $y_{ij} \in Ber(\frac{1}{2})$ se tiene la conjetura de que

$$p_n = \mathbb{P}(det(A_{n\times n}) = 0) = \left(\frac{1}{2} + o(1)\right)^n.$$

En 2010 se demostró que

$$\mathbb{P}(det(A_{n\times n}) = 0) = \left(\frac{1}{\sqrt{2}} + o(1)\right)^n$$

Un poco de historia:

1966 Erdös propone el problema al estudiar los signos de las permutaciones de los determinantes.

1967 Kólmos observa que $p_n = o(n)$.

1968 Kólmos la distribucuoón no-degenerada i.i.d, $p_n^F = O_F(1)$. Con esto surge que si $y_1, y_2, \ldots$ i.i.d. y $S_n = \sum_{i=1}^n \alpha_i y_i$, $\alpha_i \in \mathbb{R}\backslash\{0\}$ entonces, ¿cuál es $\sup_{x\in\mathbb{R}} \mathbb{P}(S_n = x)$?

En el caso $y_i \sim Ber(\frac{1}{2})$, $\alpha_i = 1\ \forall i$ se tiene que

$$\sup_{x\in\mathbb{R}} \mathbb{P}(S_n = x) = \binom{n}{\lfloor\frac{n}{2}\rfloor}\left(\frac{1}{2}\right)^n \equiv c\frac{1}{\sqrt{n}}.$$

¿Qué sucede si α_i son todos diferentes?.

Erdös estudió el siguiente problema. Para $y_i \sim Ber(\frac{1}{2})$, $\mathbb{P}(S_n \in I) \leq c\frac{1}{\sqrt{n}}$ donde $S_n \in I$. ¿Cuántas raíces comunes tienen n polinomios aleatorios?

Este problema lo propusieron Littlewood-Offord.

Szemeredi-Sarkozy probaron que

$$\sup_{x\in\mathbb{R}} \mathbb{P}\left(\sum_{i=1}^n \alpha_i y_i = x\right) = O(n^{-\frac{3}{2}})$$

Revisar Small Ball Probability.

Una matriz circulante es de la forma

$$C_n = \begin{pmatrix} y_0 & y_1 & \cdots & y_{n-1} \\ y_{n-1} & y_0 & \cdots & y_{n-2} \\ \vdots & \vdots & \ddots & \vdots \\ y_1 & y_2 & \cdots & y_0 \end{pmatrix}.$$

Observemos que sólo está definida por su primer renglón. Estas matrices aproximan a una matriz de Toeplitz. C_n es siempre diagonalizable en $\mathbb{C}$, es decir, $\exists F_n$ tal que $C_n = F_n^* D_n F_n$ donde F_n es la matriz de Fourier unitaria,

$$(F_n)_{kj} = (\omega_n^{(k-1)(j-1)})_{kj}$$

donde ω_n es la n-ésima raíz de la unidad, $\omega_n = e^{\frac{2\pi}{n}i}$ y D_n es diagonal, $(D_n)_{jj} = G_n(\omega_n^j)$ son los valores propios donde

$$G_n(z) = y_0 + y_i z + \cdots + y_{n-1} z^{n-1}.$$

Los valores singulares de la matriz C_n son $|G_n(\omega_n^k)|$ para $k = 0, \ldots, n-1$. El valor singular mínimo, S_n, de una matriz circulante es $S_n(C_n) = \min_{0\leq k\leq n-1} |G_n(\omega_n^k)|$.

Nota: Cualquier matriz $M_{n\times n}$ en $\mathbb{C}$ se puede descomponer de la siguiente manera:

$$M_n = \prod_{i=1}^{N(M_n)} N_i,$$

donde N_i es diagonal o circulante.

Mundo estocástico

Teorema 1.3. Sean $y_0, y_1, \ldots, y_{n-1}$ v.a.i.i.d,

$$\mathbb{P}(S_n(C_n) = 0) = \mathbb{P}\left(\min_{0\leq k\leq n-1} |G_n(\omega_n^k)| = 0\right) \leq \frac{c}{\sqrt{n}}.$$

La desigualdad se puede encontrar [Mec09]. Veamos la solución para el caso en que $\mathbb{P}(Y_0 \in \mathbb{Z}) = 1$. Para ello, necesitamos el siguiente resultado de Halász.

Teorema 1.4 (Halász, 1977)**.** Sean $a_1, a_2, \ldots, s_n \in \mathbb{R}^2$ todos distintos, $y_1, \ldots, y_n$ Bernoulli(-1,1) independientes entonces

$$\sup_{x\in\mathbb{R}} \mathbb{P}\left(\sum_{k=1}^{n} a_k y_k = x\right) = O(n^{-3}).$$

Ahora,

$$\mathbb{P}\left(\min_{0\leq k\leq n-1} |G_n(\omega_n^k)| = 0\right) \leq \mathbb{P}\left(|G_n(1)| = 0\right) + \sum_{1\leq k\leq n-1} \mathbb{P}\left(|G_n(\omega_n^k)| = 0\right),$$

pero

$$\mathbb{P}\left(|G_n(1)| = 0\right) = \mathbb{P}\left(\sum_{0\leq l\leq n-1} y_l \omega_n^{lk} = 0\right).$$

Para aplicar el resultado de Halász, necesitamos que para $k = 1, \ldots, n-1$, $\{\omega_n^{lk} : l = 0, \ldots, n-1\}$ son todos diferentes. Obsevemos que

$$|\{\omega_n^{lk} : l = 0, \ldots, n-1\}| \geq \phi(n) \geq \frac{n}{\log(\log(n))},$$

donde ϕ es la función Phi de Euler (Euler totient).

Observación 1.8. 1. Recordemos la función de concentración de Lévy

$$Q(X;\lambda) = \sup_{x\in\mathbb{R}} \mathbb{P}(x \leq X \leq x+\lambda), \qquad \forall\lambda \geq 0.$$

2. si X, Y son independientes, entonces

$$Q(X+Y;\lambda) \leq \min\{Q(X;\lambda), Q(Y;\lambda\}, \qquad \forall\lambda \geq 0.$$

Si usamos la función Phi de Euler y la función de concentración de Lévy, tenemos que

$$\mathbb{P}\left(\sum_{0\leq l\leq n-1} y_l\omega_n^{lk}=0\right)\leq \mathbb{P}\left(\sum_{l\in I} y_l\omega_n^{lk}=0\right)=O(n^{-3}),$$

donde $|I|=\phi(n)$. Así, para n impar

$$\mathbb{P}(S_n(C_n)=0)=O(n^{-2})$$

Descomposición Bernoulli. Dada X v.a. y $p\in(0,1)$ tenemos que $X=Y+W\eta$ con $(Y,W)\perp\eta$, $W>0$ y $\eta\sim Ber(p)$.

Sea X Bernoulli(-1,1) y Y Bernoulli, $X=LY-1$. Observemos que si $\mathbb{P}(y\in\mathbb{Z})=1$, y

$$\mathbb{P}(L=0)=\mathbb{P}(L=1)=\frac{1}{2}$$

$$\mathbb{P}(\eta=1)=1-\mathbb{P}(\eta=0)=q$$

$$q=\sum_{j\in\mathbb{Z}}\min\{\mathbb{P}(y=j),\mathbb{P}(y=j+1)\},$$

entonces $\exists Y$ tal que $y\overset{d}{=}Y+\eta L$ con $(Y,\eta)\perp L$. Usando lo anterior, tenemos que

$$\mathbb{P}(S_n(C_n)=0)=\mathbb{P}\left(\sum_{k=0}^{n-1}y_k=0\right)+O(n^{-2}).$$

Si $\mathbb{P}(y_0\in\mathbb{Z})=1$, para $k=1,\ldots,n-1$

$$\mathbb{P}\left(\sum_{0\leq l\leq n} y_l\omega_n^{lk}=0\right)=\mathbb{P}\left(\sum_{0\leq l\leq n}(Y_l+\eta_l L_l)\omega_n^{lk}=0\right)$$

$$=\mathbb{P}\left(\sum_{0\leq l\leq n}\eta_l(2L_l-1)\omega_n^{lk}=-2\sum_{0\leq l\leq n}Y_l-\sum_{0\leq l\leq n}\eta_l\omega_n^{lk}\right).$$

Consideramos el evento A_n en el que la cantidad de l tales que $\eta_l=1$ es mayor o igual $\frac{n}{2}$ y entonces tenemos que lo anterior es menor o igual a

$$\mathbb{P}\left(\sum_{0\leq l\leq n}\eta_l(2L_l-1)\omega_n^{lk}=-2\sum_{0\leq l\leq n}Y_l-\sum_{0\leq l\leq n}\eta_l\omega_n^{lk},A_n\right)+\mathbb{P}(A_n^C),$$

lo segundo decae exponencialmente y luego hay que condicionar sobre $(Y,\eta)\perp L$ para llegar a que es menor a $O(n^{-2})$.

Ahora, el problema $\forall\varepsilon>0$, $\exists c$ tal que

$$\mathbb{P}(S_n(C_n)\leq\varepsilon n^{-\frac{1}{2}})\leq C_F\varepsilon.$$

v.a. subgaussianas, esto implica que las raíces del polinomio estarían en una franja alrededor del círculo unitario.

1.4. Propiedad de Invarianza

En muchos de los ejemplos que trabajaremos se usarán clases de matrices que tienen una propiedad muy conveniente en el sentido de los cálculos y generalidad de resultados, esa propiedad será la de ser invariantes en algún sentido. En esta sección definimos algunos tipos de invarianza.

Definición 1.7.

1. Una matriz aleatoria $X \in \mathbb{M}_{d\times e}(\mathbb{R})$ es invariante bajo transformaciones ortogonales por la izquierda [derecha] si $\forall\ O \in \mathbb{O}(d)$ $[\forall O \in \mathbb{O}(e)]$, $OX \stackrel{d}{=} X$ $[XO \stackrel{d}{=} X]$.
2. Una matriz aleatoria $X \in \mathbb{M}_{d\times e}(\mathbb{C})$ es invariante bajo transformaciones unitarias por la izquierda [derecha] si $\forall\ U \in \mathbb{U}(d)$ $[\forall U \in \mathbb{U}(e)]$, $UX \stackrel{d}{=} X$ $[XU \stackrel{d}{=} X]$.

Lema 1.3. Sea $X \in \mathbb{M}_{d\times e}(\mathbb{R})$ con distribución invariante bajo transformaciones ortogonales por la izquierda (análogamente para unitarias y derecha). Sea $h : \mathbb{M}_{d\times e}(\mathbb{R}) \to \mathbb{M}_{d'\times e'}(\mathbb{R})$ una función $\mathbb{B}(\mathbb{M}_{d\times e}) \setminus \mathbb{B}(\mathbb{M}_{d'\times e'})$-medible. Entonces

$$h(OX) \stackrel{d}{=} h(X).$$

Demostración.
Sea $A \in \mathcal{B}(\mathbb{M}_{d'\times e'})$ y sea $O \in O(d)$.

$$\mathbb{P}(h(OX) \in A) = \mathbb{P}(OX \in h^{-1}(A)) \stackrel{\text{(invarianza)}}{=} \mathbb{P}(X \in h^{-1}(A)) = \mathbb{P}(h(X) \in A).$$

Por lo tanto, $h(OX) \stackrel{d}{=} h(X)$. ■

Definición 1.8. Sea X matriz aleatoria en $\mathbb{M}_d(\mathbb{R})$ $(X \in \mathbb{M}_d(\mathbb{R}))$. Se dice que X tiene distribución invariante bajo conjugaciones ortogonales (unitarias) si $\forall O \in \mathbb{O}(d)$ $(\forall U \in \mathbb{U}(d))$ no aleatorio, $OXO^t \stackrel{d}{=} X$, $(UXU^* \stackrel{d}{=} X)$.

Lema 1.4. Sea $X \in \mathbb{M}_{d\times e}(\mathbb{R})$ $(\mathbb{M}_{d\times e})$ con distribución invariante bajo conjugaciones ortogonales (unitarias) y $h : \mathbb{M}_{d\times e}(\mathbb{R}) \to \mathbb{M}_{d'\times e'}(\mathbb{R})$ una función $\mathbb{B}(\mathbb{M}_{d\times e}) \setminus \mathbb{B}(\mathbb{M}_{d'\times e'})$-medible. Entonces para todo $O \in \mathbb{O}(d)$ $(U \in \mathbb{U}(d))$ tenemos

$$h(OXO^t) \stackrel{d}{=} h(X). \quad (h(UXU^*) \stackrel{d}{=} h(X)).$$

Observación 1.9. El concepto de invarianza bajo transformaciones ortogonales $OX \stackrel{d}{=} X$, es una extensión del concepto de variable aleatoria simétrica $(-X \stackrel{d}{=} X)$. No así el de invarianza bajo conjugaciones ortogonales.

Lema 1.5. Sea $X \in \mathbb{S}_d$ $(X \in \mathbb{H}_d)$ invariantes bajo conjugaciones ortogonales (unitarias). Entonces

$$c_X(T) = \mathbb{E}e^{itr(TX)} = \mathbb{E}e^{iT_E diag(X)^t} \qquad \forall T \in \mathbb{S}_d\ (T \in \mathbb{H}_d),$$

donde $T_E = (\lambda_1(T), \ldots, \lambda_d(T))$ es el vector de valores propios de T. Es decir,

$$c_X(T) = \mathbb{E}e^{i\sum_{j=1}^d \lambda_j(T)X_{jj}} = c_{(X_{11},\ldots,X_{dd})}(\lambda_1(T), \ldots, \lambda_d(T)).$$

Demostración.

Sea $T \in \mathbb{S}_d$ y $O \in \mathbb{O}(d)$ tal que

$$T = O^t Diag(T_E) O = O^t \begin{pmatrix} \lambda_1(T) & & 0 \\ & \ddots & \\ 0 & & \lambda_d(T) \end{pmatrix} O.$$

Por lo tanto,

$$\begin{aligned} \mathbb{E}e^{iTr(TX)} &= \mathbb{E}e^{iTr(O^t Diag(T_E)OX)} = \mathbb{E}e^{iTr(Diag(T_E)OXO^t)} \\ &\overset{\text{(invarianza)}}{=} \mathbb{E}e^{iTr(Diag(T_E)X)} = \mathbb{E}e^{i\sum_{j=1}^d \lambda_j(T)X_{jj}}, \end{aligned}$$

que es lo que queríamos demostrar. ■

1.5. Ensambles de Matrices

1.5.1. Ensambles Gaussianos

Sea Z matriz $d \times d$ Ginibre gaussiana en $\mathbb{M}_d(\mathbb{R})$. Definimos $G \in \mathbb{S}_d$ como

$$G = \frac{1}{\sqrt{2}}(Z + Z^t). \tag{1.4}$$

Si $Z \in \mathbb{M}_d(\mathbb{C})$, definimos $G \in \mathbb{H}_d$ como

$$G = \frac{1}{\sqrt{2}}(Z + Z^*). \tag{1.5}$$

Proposición 1.3. $G \in \mathbb{S}_d$ como en 1.4 es invariante bajo conjugaciones ortogonales. Análogamente $G \in \mathbb{H}_d$ como en 1.5 es invariante bajo conjugaciones unitarias.

Demostración.

Caso ortogonal (unitario es análogo). Queremos ver que $\forall O \in \mathbb{O}(d)$ se tiene que $O^t G O \overset{d}{=} G$. En efecto,

$$O^t G O = \frac{1}{\sqrt{2}}(O^t Z O + O^t Z^t O) \overset{d}{=} \frac{1}{\sqrt{2}}(ZO + O^t Z^t) = \frac{1}{\sqrt{2}}(Z + Z^t) = G.$$

■

Observación 1.10. Analicemos aspectos distribucionales de ambos casos de G ((1.4) y (1.5)).

i) Consideremos primero el caso en el que $G \in \mathbb{H}_d$; como $Z_{ij} = Re(Z_{ij}) + iIm(Z_{ij})$, con

$Re(Z_{ij}), Im(Z_{ij})$ independientes, ambas con distribución $N(0,\frac{1}{2})$. Tenemos entonces que

$$G_{ii} = \frac{1}{\sqrt{2}}[Z_{ii} + \bar{Z}_{ii}] = \frac{1}{\sqrt{2}}[2Re(Z_{ii})] = \sqrt{2}Re(Z_{ii}),$$

y entonces

$$\mathbb{E}(G_{ii}) = 0, \quad \mathbb{V}(G_{ii}) = \mathbb{V}(\sqrt{2}Re(Z_{ii})) = 2\mathbb{V}(Re(Z_{ii})) = 2\cdot\frac{1}{2} = 1.$$

Para el caso $i \neq j$, usaremos que las variables G_{ij} son centradas (fácil de ver) y el hecho de que si Y variable aleatoria compleja centrada, entonces $\mathbb{V}(Y) = \mathbb{E}(|Y|^2)$. Así pues, como

$$|G_{ij}|^2 = G_{ij}\bar{G}_{ji} = \frac{1}{2}[Z_{ij}\bar{Z}_{ji} + Z_{ij}\bar{Z}_{ij} + \bar{Z}_{ij}\bar{Z}_{ji} + Z_{ij}\bar{Z}_{ji}],$$

entonces, por independencia:

$$\begin{aligned}
\mathbb{V}(G_{ij}) &= \mathbb{E}[|G_{ij}|^2] = \frac{1}{2}\mathbb{E}[|Z_{ij}|^2 + |\bar{Z}_{ji}|^2] = \frac{1}{2}\mathbb{E}[2|Z_{ij}|^2] \\
&= \mathbb{E}[[Re(Z_{ij})]^2 + [Im(Z_{ij})]^2] = \mathbb{V}(Re(Z_{ij})) + \mathbb{V}(Im(Z_{ij})) \\
&= \frac{1}{2} + \frac{1}{2} = 1
\end{aligned}$$

Concluimos que $G_{ij} \sim N(0,1)$ para todo i,j. Un simple cálculo nos da también la función característica en el caso unitario:

$$c_G(T) = e^{-\frac{1}{4}Tr(T^2)} \qquad \forall T \in \mathbb{H}_d.$$

ii) Para el caso en que $G \in \mathbb{S}_d$ tenemos resultados diferentes ya que

$$G_{ij} = \begin{cases} \sqrt{2}Z_{ii}, \ i = j \\ \frac{1}{\sqrt{2}}(Z_{ij} + Z_{ji}), \ i \neq j \end{cases}.$$

y por tanto $\mathbb{V}G_{ii} = 2\mathbb{V}(Z_{ii}) = 2$ y para $i \neq j$, $\mathbb{V}G_{ij} = 1$. Concluimos que $(G_{ij})_{1\leq i\leq j\leq d}$ son variables independientes con distribución:

$$G_{ij} \sim \begin{cases} N(0,2), \ i = j \\ N(0,1), \ i \neq j \end{cases}.$$

Sobre la función característica en este caso, sea $T \in \mathbb{S}_d$, entonces

$$c_G(T) = \mathbb{E}e^{i\sum_{j=1}^{d}\lambda_j(T)(G_1)_{jj}} = \prod_{j=1}^{d} e^{-\frac{2}{2}\lambda_j^2(T)} = e^{-\sum_{j=1}^{d}\lambda_j^2(T)} = e^{-Tr(T^2)}.$$

Por tanto

$$c_G(T) = e^{-Tr(T^2)} \qquad \forall T \in \mathbb{S}_d.$$

Definición 1.9. Un *ensamble* $(X_n)_{n\geq 1}$ es una sucesión de matrices aleatorias tal que para todo $n \geq 1$, X_n es una matriz $n \times n$.

Definición 1.10. 1) Un ensamble $G = (G^d)_{d\geq 1}$ se dice *Gaussiano Ortogonal* y se abrevia GOE si $\forall d \geq 1$, $G^d = (G^d_{ij})_{i,j=1,\ldots,d}$ es tal que $\{G^d_{ij} : 1 \leq i \leq j \leq d\}$ son variables aleatorias independientes, $G^d \in \mathbb{S}_d$ y

$$G^d_{ii} \sim N(0,2)$$

$$G^d_{ij} \sim N(0,1) \qquad \forall i \neq j.$$

Es decir, G^d es como G definida en (1.4).

2) Un ensamble $G = (G^d)_{d\geq 1}$ se dice *Gaussiano Unitario)* y se abrevia GUE si $\forall d \geq 1$, $G^d = (G^d_{ij})_{i,j=1,\ldots,d}$ es tal que $\{G^d_{ij} : 1 \leq i \leq j \leq d\}$ son variables aleatorias independientes, $G^d \in \mathbb{H}_d$ y $G^d_{ij} \sim N(0,1) \qquad \forall i,j$. Es decir, G^d es como G en (1.5).

Recordemos de la definición 1.4 que una matriz de Wigner $X \in \mathbb{S}_d$ (o $X \in \mathbb{H}_d$) cumple que $\{X^d_{ij} : 1 \leq i \leq j \leq d\}$ son v.a. independientes. Es claro de la definición que las matrices de los ensambles GUE y GOE son de Wigner; lo interesante es que también tenemos un recíproco, lo presentamos a continuación.

Teorema 1.5 (Proyecto del 2017)**.** Sea $X \in \mathbb{S}_d$ matriz aleatoria de Wigner, no diagonal e invariante bajo conjugaciones ortogonales. Entonces X es $GOE(d)$.

Las siguientes expresiones serán muy útiles para futuros cálculos.

Observación 1.11. Para ambos casos ($G \in \mathbb{S}_d$ o $G \in \mathbb{H}_d$, $\mathbb{P}(det(G) \neq 0) = 1$. Las funciones de densidad son

$$f_G(z) = \left(\frac{1}{2}\right)^d \left(\frac{1}{\sqrt{2\pi}}\right)^{\frac{d(d+1)}{2}} e^{-Tr(z^2)} \qquad z \in \mathbb{S}_d, \quad (G \in \mathbb{S}_d),$$

$$f_G(z) = \left(\frac{1}{2}\right)^{\frac{d}{2}} \left(\frac{1}{\pi}\right)^{\frac{d^2}{2}} e^{-\frac{1}{2}Tr(z^2)} \qquad z \in \mathbb{H}_d, \quad (G \in \mathbb{H}_d).$$

1.5.2. Ensamble Wishart

Definimos ahora otro ensamble de matrices, el cuál será de vital importancia en la parte de aplicaciones.

Definición 1.11. Sea $Z = \{Z_{ij}\} \in M_{d\times e}(\mathbb{R})$ con $Z_{ij} \sim N(0,1)$ independientes para $1 \leq i \leq d$, $1 \leq j \leq e$, $e \leq d$. Definimos por $W := Z^t Z$ matriz $e \times e$, observemos que $W \in \mathbb{S}^+_e$. A la matriz W se le conoce como la **matriz de Wishart**.

Nota 1.2. También trabajaremos con el caso complejo de la Wishart:
En este caso tomamos $Z = \{Z_{ij}\} \in M_{d\times e}(\mathbb{C})$ con Z_{ij} gaussianas complejas independientes y $W := Z^* Z$ es el caso complejo de la matriz de Wishart, observemos que $W \in \mathbb{H}^+_e$.

Observación 1.12. 1. En el caso $e = 1$, W es una χ^2 con d grados de libertad. Podemos pensar así a la matriz de Wishart como el análogo matricial de tal distribución.

2. La matriz de Wishart (1928) es la primera matriz aleatoria. El caso 2×2 lo estudió Fisher.

3. La matriz de $d \times d$: $\tilde{W} = ZZ^t$ se conoce como la matriz anti-Wishart (llamada así por físicos) o la Wishart singular ya que no tiene rango completo.

4. Equivalentemente se puede definir la distribución de Wishart como la distribución de $A = \sum_{j=1}^{n} x_j x_j^T$, donde $x_1, \ldots, x_n$ v.a.i.i.d $N_d(0, I)$ y usando cambio de variable podemos llegar a que la densidad es:

$$f(A) \equiv W_d(A; n, I) = \frac{e^{-\frac{1}{2}Tr(A)} \; |A|^{\frac{n-d-1}{2}}}{2^{\frac{nd}{2}} \; \Gamma_d(\frac{n}{2})}.$$

5. Variando la definición podemos obtener la distribución de Wishart con matriz de covarianzas Σ como la distribución de $A = \sum_{j=1}^{n} x_j x_j^T$, donde $x_1, \ldots, x_n$ v.a.i.i.d $N_d(0, \Sigma)$ y de nuevo por cambio de variable tenemos la densidad es:

$$f(A) \equiv W_d(A; n, \Sigma) = \frac{e^{-\frac{1}{2}Tr(\Sigma^{-1}A)} \; |A|^{\frac{n-d-1}{2}}}{2^{\frac{nd}{2}} \; \Gamma_d(\frac{n}{2}) \; |\Sigma|^{\frac{n}{2}}}.$$

Nota 1.3. En la definición análoga de la distribución Wishart de la observación anterior, debemos tener cuidado si las observaciones son vectores columna o renglones; diferentes libros usan diferente notación.

Enlistamos en la siguiente proposición algunas propiedades de la distribución Wishart.

Proposición 1.4. 1. La Matriz Wishart es no singular con probabilidad 1, es decir $\mathbb{P}(\det(W) \neq 0) = 1$.

2. W es invariante bajo conjugaciones ortogonales (unitarias en caso complejo).

3. Transformada de Laplace:

$$\mathbb{E}e^{-Tr(TW)} = \det(I + 2T)^{-\frac{d}{2}} \qquad T \geq \frac{1}{2}I, \; T \in \mathbb{S}_e^+,$$

que recordando de la primera sección $A \geq B$ si $A - B$ es definida positiva.

4. Z y W son matices aleatorias independientes. Es decir,

$$\mathbb{P}(Z \in A, W \in B) = \mathbb{P}(Z \in A)\mathbb{P}(W \in B) \qquad \forall A \in \mathcal{B}(M_{d\times e}), B \in \mathcal{B}(S_e^+).$$

Se concluye que la función característica conjunta de Z y W es el producto de las funciones características de Z y W. Es decir,

$$\mathbb{E}e^{iTr(T^tZ)}e^{iTr(SW)} = \mathbb{E}e^{iTr(T^tZ)}\mathbb{E}e^{iTr(SW)} \qquad \forall T \in M_{d\times e}, S \in S_e^+.$$

Demostración.

1. Observemos primero que como las entradas de Z son independientes y con distribución continua, por lo visto en la subsección 1.3.1 concluimos $\mathbb{P}(rango(Z) = e) = 1$ (rango completo). Por lo tanto $\mathbb{P}(\det(W) \neq 0) = 1$. En el caso de la Wishart con matriz de covarianzas Σ, también tenemos la no singularidad porque es de la forma $W\Sigma$ y W no singular y Σ positiva definida.

2. Sea $O \in \mathbb{O}(e)$, entonces
$$OWO^t = OZ^tZO^t = (ZO^t)^tZO^t = Z^tZ = W.$$

3. Como $T \in \mathbb{S}_e^+$, podemos diagonalizarla con una matriz ortogonal O, de la siguiente manera:
$$T = O(diag(t_i))O^t.$$
Además, sabemos que W es invariante bajo conjugaciones ortogonales. Por lo tanto,
$$\begin{aligned}\mathbb{E}e^{-Tr(TW)} &= \mathbb{E}e^{-Tr(O(diag(t_i))O^tW)} = \mathbb{E}e^{-Tr((diag(t_i))O^tWO)} \\ &= \mathbb{E}e^{-Tr((diag(t_i))W)} = \mathbb{E}e^{-\sum_{i=1}^{e} t_iW_{ii}}.\end{aligned}$$
Recordemos que $W = Z^tZ$, donde $Z = Z_{d\times e} = \{Z_{ij}\}$ con $Z_{ij} \sim N(0,1)$ para $1 \leq i \leq d$, $1 \leq j \leq e$, $e \leq d$. Entonces $W_{ii} = \sum_{j=1}^{d} Z_{ji}^2$ se distribuye como una χ^2 con d grados de libertad, que sabemos tiene FGM $(1-2t)^{-\frac{d}{2}}$ para $t < \frac{1}{2}$. Por otro lado, como $Z_{i_1j_1} \perp Z_{i_1j_2}$ si $i_1 \neq i_2$, entonces $W_{ii} \perp W_{jj}$ si $i \neq j$. Entonces tenemos que
$$\mathbb{E}e^{-\sum_{i=1}^{e} t_iW_{ii}} = \prod_{i=1}^{e}\mathbb{E}e^{-t_iW_{ii}} = \prod_{i=1}^{e}(1+2t_i)^{-\frac{d}{2}} = \left(\prod_{i=1}^{e}(1+2t_i)\right)^{-\frac{d}{2}}$$
$$= (\det(I + 2(diag(t_i))))^{-\frac{d}{2}} = \big(\det(OIO^t + 2O(diag(t_i))O^t)\big)^{-\frac{d}{2}} = (\det(I+2T))^{-\frac{d}{2}}$$
donde necesitamos la condición de que $-t_{ii} < \frac{1}{2}$ equivalente a $-\frac{1}{2} < t_{ii}$ para poder usar la FGM de cada χ^2. Y esa condición se transforma a que $I < 2T$ o $\frac{I}{2} < T$. Por lo tanto, concluimos que
$$\mathbb{E}e^{-Tr(TW)} = \det(I+2T)^{-\frac{d}{2}} \qquad T > \frac{I}{2}, T \in \mathbb{S}_e^+.$$

■

1.5.3. Ensambles Ortogonal y Unitario

Establezcamos algo de notación que facilitará la exposición de resultados.

Notación 1.1. Para $A \in \mathbb{S}_e^+$ podemos encontrar $O \in \mathbb{O}(e)$ tal que
$$A = O\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_d\end{pmatrix}O^t,$$

y dada esa descomposición de A, podemos definir

$$A^{\frac{1}{2}} = O \begin{pmatrix} \lambda_1^{\frac{1}{2}} & & 0 \\ & \ddots & \\ 0 & & \lambda_d^{\frac{1}{2}} \end{pmatrix} O^t.$$

En general dada la descomposición podemos definir:

$$F(A) = O \begin{pmatrix} F(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & F(\lambda_d) \end{pmatrix} O^t,$$

lo que llamamos el "cálculo funcional" (se definirá más adelante). Cabe decir que las relaciones anteriores existen y son únicas al fijar un orden de los eigenvalores de la matriz diagonalizada, por ejemplo $\lambda_1 \geq \ldots \geq \lambda_d$.

Consideremos $Z \sim N(0,1)$ (matriz gaussiana con $d = e = 1$). Como Z no se anula, podemos definir:

$$O = \frac{Z}{|Z|} = \frac{Z}{|Z^2|^{\frac{1}{2}}},$$

y observamos que sólo toma los valores -1 y 1, además

$$\mathbb{P}[O = 1] = \mathbb{P}[Z > 0] = \frac{1}{2} = \mathbb{P}[Z < 0] = \mathbb{P}[O = -1],$$

es decir O es Bernoulli Simétrica o Rademacher. También observemos que el grupo ortogonal de dimensión 0 es $\{-1, 1\}$ y por tanto la matriz O es ortogonal. Todo esto lo generalizamos en la siguiente observación, usando la notación establecida antes.

Observación 1.13. Ahora, dado $Z = Z_{d \times e} = (Z_{ij})$ con $Z_{ij} \sim N(0,1)$ v.a.i.i.d. tomemos la matriz $W = Z^t Z$. La matriz W es claramente diagonalizable (simétrica real) y positiva semidefinida pero por lo visto antes también es del tipo M y por tanto W^{-1} existe casi seguramente, así que concluimos que $W \in \mathbb{S}_e^+$. Por lo establecido antes $W^{\frac{1}{2}} = (Z^t Z)^{\frac{1}{2}}$ existe y es no singular con probabilidad 1. Además $(Z^t Z)^{-\frac{1}{2}}$ también existe y es no singular (ya que todos los eigenvalores son estrictamente positivos no afecta tomar el inverso y raíz).

Vamos ahora a definir y a estudiar dos matrices V_R y V_I que nos permitirán concluir cosas interesantes del grupo ortogonal y unitario.

Proposición 1.5. Sea $e \leq d$ y sean $V_R = (Z^t Z)^{-\frac{1}{2}} Z^t$ una matriz de $e \times d$ y $V_I = Z(Z^t Z)^{-\frac{1}{2}}$ una matriz de $d \times e$. Entonces

1. $\mathbb{P}(V_R V_R^t = I_e) = 1, \quad \mathbb{P}(V_I^t V_I = I_e) = 1 \quad$ y $\quad \mathbb{P}(V_R^t V_R \in \mathbb{S}_d^+, V_I V_I^t \in \mathbb{S}_d^+) = 1.$
2. Si $d = e$, entonces $(Z^t Z)^{-1} = Z^{-1}(Z^{-1})^t$ y tenemos que

$$\mathbb{P}(V_R^t V_R = V_R V_R^t = I_d) = 1,$$

$$\mathbb{P}(V_I^t V_I = V_I V_I^t = I_d) = 1.$$

3. Si $d = e$, entonces $\mathbb{P}(V_R \in \mathbb{O}(d)) = 1$.

4. Si $d = e$, entonces $\forall O \in \mathbb{O}(d)$ tenemos que

$$\mathbb{P}(OV_R \in A) = \mathbb{P}(V_R \in A) \qquad \forall A \in \mathcal{B}(\mathbb{O}(d)),$$

$$\mathbb{P}(V_R O \in A) = \mathbb{P}(V_R \in A) \qquad \forall A \in \mathcal{B}(\mathbb{O}(d)).$$

Observación 1.14.

1. A la distribución de V_R (en el caso $d = e$) se le conoce como la distribución de Haar (o distribución uniforme).
2. Lo anterior da una demostración de la existencia de la distribución de Haar en $\mathbb{O}(d)$ ($\mathbb{U}(d)$) usando matrices aleatorias. (Las construimos así, el punto fino fue la no singularidad).
3. En el caso $d = e = 1$, lo único que importó de la distribución normal fue que fuera simétrica.
4. Se puede probar que también se puede hacer la construcción de Haar si comenzamos con una matriz X ($d \times d$) cuya distribución es invariante bajo transformaciones por la izquierda y es no singular, por ejemplo, $X = \sigma Z$ con $Z = (Z_{ij})_{1 \leq i,j \leq d}$ independientes $N(0,1)$ y σ cualquier variable aleatoria no negativa.
5. Este método tiene la gran ventaja de que da una forma de simular matrices aleatorias con distribución de Haar.

1.6. Aspectos Distribucionales

Lo presentado en esta última sección es lo referido a *Teoría Clásica* de matrices aleatorias. Veremos aquí el análogo matricial de varias de las distribuciones muestrales, la distribución normal multivariada, entre otras que servirán de referencia en lo posterior de las notas.

Definición 1.12. El vector aleatorio X de dimensión m tiene *distribución normal multivariada* con media $\mu \in \mathbb{R}^m$ y matriz de covarianzas $\Sigma \in \mathbb{S}_m^+$, y escribimos $X \sim N_m(\mu, \Sigma)$, si tiene densidad f_X, donde $\forall x \in \mathbb{R}^m$:

$$f_X(x) = \left(\frac{1}{\sqrt{2\pi}}\right)^m \frac{1}{(\det(\Sigma))^{1/2}} \exp\left(-\frac{1}{2}(X-\mu)^T \Sigma^{-1}(X-\mu)\right).$$

Es un ejercicio análogo al caso univariado probar la siguiente proposición.

Proposición 1.6. Si $X = (X_1, \ldots, X_m)$.

a) $\mu = \mathbb{E}[X] = (\mathbb{E}(X_1), \ldots, \mathbb{E}(X_m))^T$ y $\Sigma = \mathbb{E}[(X-\mu)^T(X-\mu)] = (Cov(X_i, X_j))$.

b) La función característica de X, es $C_X(T) = \exp(iT^t\mu - \frac{1}{2}T^t\Sigma T)$.

c) Si $X \sim N_m(0, I_m)$ y $\Sigma = \Sigma_{m\times m} > 0$, entonces

$$Y = \Sigma^{1/2} X \sim N_m(0, \Sigma),$$

y $Y + \mu \sim N_m(\mu, \Sigma)$. Viceversa, si $Y \sim N_m(0, \Sigma)$ entonces $\Sigma^{-1/2} Y \sim N_m(0, I_m)$.

Definición 1.13 (Análogo Matricial). X matriz aleatoria en $\mathbb{M}_{d\times e}(\mathbb{R})$ tiene distribución normal, y lo denotamos $N_{d\times e}(\mu, C\otimes D)$, si $\mu \in M_{d\times e}(\mathbb{R})$ no aleatoria, $C = C_{d\times d} > 0,\ D = D_{e\times e} > 0$, $C \otimes D$ es el producto de Kronecker (definido en la sección 1.1) y si $vec(X)$ es la matriz X vista como vector de $\mathbb{R}^{de}$, entonces

$$vec(X) \sim N_{de}(vec(\mu), C\otimes D),$$

es decir es vector gaussiano con esas dimensiones y parámetros.

A continuación se enlistan hechos de las matrices con distribución normal.

a) La densidad de X en $\mathbb{M}_{d\times e}(\mathbb{R}) \cong \mathbb{R}^{de}$ es

$$f_X(x) = \left(\frac{1}{2\pi}\right)^{\frac{de}{2}} \det(C)^{-\frac{d}{2}} \det(D)^{-\frac{e}{2}} \exp\left(-\frac{1}{2}Tr\left[C^{-1}(X-\mu)D^{-1}(X-\mu)^T\right]\right).$$

b) La media es $\mu = \mathbb{E}[X] := (\mathbb{E}(X_{ij}))$, siempre que $X = (X_{ij})$.

c) Dos variantes de la varianza:

$$D = \mathbb{E}(X-\mu)^T(X-\mu)Tr(C)$$
$$C = \mathbb{E}(X-\mu)(X-\mu)^T Tr(D)$$

Por lo tanto

- Si $C = I_d$ entonces los renglones de X son independientes.
- Si $D = I_e$ entonces las columnas de X son independientes.

d) La función característica de X es

$$C_X(T) = \exp\left(iTr(T^t\mu) - \frac{1}{2}Tr(CTDT^t)\right), \quad T \in \mathbb{M}_{d\times e}(\mathbb{R}).$$

e) Si $C = I_d$ y $D = I_e$, X es matriz de Ginibre rectangular normal.

f) De especial interés en estadística multivariada es el caso $p = e$, $n = d$, $C = I_n$ y $D = \Sigma$ en cuyo caso $X = (\vec{x_1}, \ldots, \vec{x_n})$, con $\vec{x_1}, \ldots, \vec{x_n}$ vectores aleatorios independientes con la misma distribución $N_p(\mu, \Sigma)$.

   El interés es estimar μ y Σ, la media y matriz de covarianza. Típicamente C tiene que ver con la muestra (diseño del experimento) y D tiene que ver con las propiedades de interés.

g) El estimador de máxima verosimilitud de μ y Σ son

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} \vec{X}_i \ \text{ y } \hat{S} = \frac{1}{n}A$$

donde

$$A = \sum_{i=1}^{n} (\vec{X}_i - \bar{X})(\vec{X}_i - \bar{X})^T.$$

Se tiene que $\mathbb{E}[\bar{X}] = \mu$ y $\mathbb{E}[\hat{S}] = \frac{n-1}{n}\Sigma$ y $\mathbb{E}[S] = \Sigma$, donde $S = \frac{1}{n-1}A$.

h) Como en el caso univariado $\bar{X}$ y A son independientes. Donde la independencia de un vector con una matriz la corroboramos con que podemos separar la función característica. Podemos probar además que $\bar{X} \sim N(\mu, I_n \otimes \Sigma)$ y $A \sim W(n-1, \Sigma)$ (distribución Wishart).

i) Si $Z \sim N(0, I_n \otimes I)$ decimos que $A = Z^T Z \in \mathbb{S}_p^+(\mathbb{R})$ tiene distribución Wishart $W(n, \Sigma)$, cuya densidad cuando $n > p - 1$:

$$f_A(B) = \frac{1}{2^{np/2}\Gamma_p(n/2)} \exp\left(-\frac{1}{2}\Sigma^{-1}B\right) \det(B)^{\frac{n-p-1}{2}}$$

donde Γ_p es la función gamma multivariada.

j) $\mathbb{E}[A] = n\Sigma$ y $\mathbb{V}(A_{ij}) = n(\sigma_{ij}\sigma_{ji} + \sigma_{ii}\sigma_{jj}) = n(\sigma_{ij}^2 + \sigma_{ii}\sigma_{jj})$.

k) La función característica de $A \sim W(n, \Sigma)$

$$C_A(T) = \det(I_p + 2iT\Sigma)^{-n/2}, \quad T > 0 \qquad = \det(I_p + 2i\Sigma^{1/2}T\Sigma^{1/2})^{-n/2}, \quad T \in \mathbb{S}_p^+.$$

l) Podemos pensar en análogos matriciales de varias distribuciones muestrales (surge de manera natural la distr. t, F, la beta, con $ZA^{-1/2}/p$ o $W_1W_2^{-1}/q$).

Jhon Wishart (1898 - 1956) nació en Montrose Escocia, asistió a la Academia de Perth y luego, en 1916 ingresó en la Universidad de Edimburgo. Allí fue profesor de matemáticas de ET Whittaker. La Primera Guerra Mundial hizo que la carrera universitaria de Wishart fuese interrumpida. Completó sus estudios universitarios en 1922, donde se graduó en matemáticas y física. Había tomado un curso de formación del profesorado en la Casa de Moray, como parte de su carrera y, después de graduarse, se trasladó a Leeds tras aceptar un puesto como profesor de matemáticas en la preparatoria West Leeds.

En 1924, después de una recomendación de Whittaker, le ofrecieron el puesto de asistente de Pearson en el University College de Londres. Wishart aprendido mucho de la estadística durante sus tres años con Pearson.
Wishart creó un laboratorio en Cambridge para sus alumnos de postgrado, donde Williams Cochran fue su alumno. Wishart tenía mucho talento para la estadística matemática y un instinto para realizar aplicaciones prácticas en diseño experimental. Posteriormente, Wishart se convirtió en el Jefe del Laboratorio de Estadística de la Facultad de Matemáticas en Cambridge.

Algunas de las publicaciones más importantes de Wishart se realizaron en el período 1928-1932 antes de que se involucrara en la docencia en Cambridge. En 1928 generalizó la distribución Chi-cuadrado que se denomina la Distribución de Wishart. También estudió las propiedades de la distribución del coeficiente de correlación múltiple que Fisher había considerado antes. También escribió numerosos artículos sobre las aplicaciones agrícolas de las estadísticas. Wishart estuvo muy involucrado con el trabajo de la Royal Statistical Society. Él fue uno de los becarios que formaron Comité Organizador de la Sección de Investigación de la Agricultura en 1933. En 1945 se convirtió en presidente de la Royal Statistical Society de la sección de investigaciones.

Wishart murió en un accidente en Acapulco, México cuando se encontraba de visita en calidad de representante de la Organización de las Naciones Unidas para la Agricultura y organizando la creación de un centro de investigación para aplicar técnicas estadísticas en la investigación agrícola.

Eugene Paul Wigner (Budapest, 1902 - Princeton, 1995) Físico norteamericano de origen húngaro. Profesor del Politécnico de Berlín desde 1930, se estableció en Estados Unidos en 1934 y a partir 1938 ejerció como profesor de física matemática en Princeton.

Los trabajos de Eugene Paul Wigner versaron sobre la física de sólidos, los núcleos atómicos y los reactores nucleares. En física nuclear, formuló el principio de la simetría de las partículas elementales o de conservación de la paridad; es muy conocida su hipótesis de que las energías potenciales de interacción entre nucleones son iguales si tienen el mismo momento angular y el mismo spin. Descubrió asimismo el efecto que lleva su nombre («efecto Wigner»), que consiste en el desplazamiento de un átomo en una red cristalina bajo la acción de un neutrón o de un ión de energía suficiente.

Eugene Wigner figuró entre los científicos que, durante la Segunda Guerra Mundial, plantearon a la administración estadounidense la necesidad de desarrollar armas atómicas, ante el temor de que estuvieran a punto de fabricarlas los alemanes, y participó en el diseño del primer reactor nuclear y de la bomba atómica. Por sus trabajos sobre los principios de la simetría en las partículas elementales recibió el premio Nobel de Física en 1963, junto a Maria Goeppert-Mayer y Hans Daniel Jensen.

Cursó estudios secundarios en Budapest en el Fasori Gimnázium luterano. Wigner fue considerado un excelente estudiante, pero no brillante. A lo largo de su vida, se refirió a su deuda con dos hombres que conoció en esa escuela. El primero fue su profesor de matemáticas, Laslo Ratz, quien reconoció que el joven Wigner tenía una excepcional habilidad en matemáticas. El segundo era un estudiante un año más joven, John von Neumann, que provenía de una familia de banqueros ricos y que fue reconocido por Ratz como un genio de las matemáticas y a quien también le brindaba clases particulares. Wigner formó una estrecha amistad con von Neumann que duró toda su vida. Cuando eran estudiantes, a menudo caminaban juntos a casa, mientras von Neumann ponía al tanto a Wigner de las maravillas de las matemáticas avanzadas.

Capítulo 2

Análisis espectral asintótico de matrices Aleatorias

Entre 1928 y 1955 se desarrolla la teoría clásica (Wishart) donde la dimensión es fija y se estudia la estadística asintótica. A partir de 1955 aparece la teoría de matrices aleatorias y se estudia cuando la dimensión tiende a infinito.

2.1. Distribución Empírica Espectral (DEE)

Sea X_N una matriz $N \times N$, no necesariamente aleatoria y sean $\lambda_1, \ldots, \lambda_N$ sus valores propios. Si X es hermitiana se define la función de distribución empírica espectral

$$F^{X_N} : \mathbb{R} \to [0, 1]$$

$$x \to \frac{1}{N} \sum_{j=1}^{N} 1_{\lambda_j \leq x}.$$

Si X_N no es hermitiana, la DEE se define como

$$F^{X_N} : \mathbb{C} \to [0, 1]$$

$$(x, y) \to \frac{1}{N} \sum_{j=1}^{N} 1_{Re(\lambda_j) \leq x, Im(\lambda_j) \leq y}.$$

En ambos casos $F^{X_N} \equiv \frac{1}{N} \sum_{j=1}^{N} \delta_{\lambda_j}$, es decir, F^{X_N} es la distribución uniforme en $\{\lambda_1, \ldots, \lambda_N\}$.

Observación 2.1.

1. No hemos pedido necesariamente que X_N sea aleatoria.

   (Nota: Esto nos da un modelo aleatorio distinto al de Kolmogorov, pues a matrices no aleatorias, les damos una distribución, es decir, que matrices con los mismos eigenvalores tendrán la misma distribución.)

2. Muchas estadísticas de interés son funcionales de la DEE.

2.2. Breviario Histórico

Esto va al último

En los dos capítulos anteriores se discutieron aspectos específicos de la teoría de matrices aleatorias; presentamos ahora un resumen a manera de breviario cultural posterior a conocer los temas de manera específica.

2.2.1. Extras del ensamble Wishart

Para A matriz aleatoria con distribución Wishart $W_p(n, \Sigma), \Sigma > 0$.

i) La función característica es

$$c_A(T) = \det(\mathrm{I}_p + 2iT\Sigma)^{-\frac{n}{2}}, T > \frac{\mathrm{I}_p}{2}, T > 0.$$

ii) La densidad en $\mathbb{R}^{p(p+1)/2}$:

$$f(A) = \frac{\det(\Sigma)^{-n/2}}{2^{np/2}\Gamma_p(\frac{n}{2})} \exp(-\frac{1}{2}Tr(\Sigma^{-1}A)) \det(A)^{(n-p-1)/2}, A > 0, n \geq p$$

$$\Gamma_p(\frac{n}{2}) = \prod_{j=1}^{p} \Gamma(a + (1-j)/2), a > (p+1)/2.$$

iii) Además $Tr(A) \sim \chi^2_{np+p(1-p)/2}$ y

$$\frac{\det(A)}{\det(\Sigma)} \sim \prod_{j=1}^{p} \chi^2_{n-j+1}, \chi^2_{n-j+1} \text{ independientes}, \quad j = 1, ..., p.$$

iv) Si $A \sim W_p(n, \Sigma)$ y $\Theta \in \mathbb{M}_p(\mathbb{R})$ es no singular, entonces

$$\Theta A \Theta^t \sim W_p(n, \Theta\Sigma\Theta^t).$$

v) $A \sim W_p(n, \Sigma)$ y $\Sigma^{-1} = \Theta^t\Theta$, entonces $\Theta A \Theta^t \sim W_p(n, \mathrm{I}_p)$.

vi) Si $A \sim W_p(n, \mathrm{I}_p)$ y $O \in O(p)$ es constante o *aleatoria independiente de* A, entonces

$$OAO^t \sim A.$$

vii) Si $A_j \sim W_p(n_j, \Sigma)$ son independientes, $j = 1, ..., k$, entonces

$$\sum_{j=1}^{k} A_j \sim W_p(\sum_{j=1}^{k} n_j, \Sigma).$$

Otros resultados:

viii) $A \sim W_p(n, \Sigma)$ y Θ matriz qxp con rango$(\Theta) = q \leq p$, entonces $\Theta A \Theta^t \sim W_q(n, \Theta\Sigma\Theta^t)$.

ix) $A \sim W_p(n, \Sigma)$ e y vector aleatorio $px1$ independiente de A y $\mathbb{P}(y \neq 0) = 1$ entonces

$$\frac{y^t A y}{y^t \Sigma y} \sim \chi_n^2 \text{ y es independiente de } y.$$

x) $A \sim W_p(n, \Sigma)$ si y sólo si $A = Z^t Z$ con $Z \sim N_{p,n}(0, \mathrm{I}_n \otimes \Sigma)$.

xi) Momentos, distribuciones marginales y condicionales a bloques.

xii) Distribución conjunta de valores propios de $A \sim W_p(n, \Sigma)$.

xiii) Densidad de A^{-1}, la Wishart inversa.

xiv) Resultados asintóticos de funciones de $A \sim W_p(n, \Sigma)$ para p fija y $n \to \infty$.

2.2.2. Más sobre el ensamble gaussiano

Si $A > 0$ es una matriz aleatoria pxp con función de densidad $f(A)$, la densidad conjunta de sus valores propios $\lambda_1 > \lambda_2 > \cdots \lambda_p$ es

$$\frac{\pi^{p^2/2}}{\Gamma_p(\frac{1}{2}p)} \prod_{i<j}^{p} (\lambda_j - \lambda_i) \int_{O(p)} f(HLH^t)\mu(dH)$$

donde $L = diag(\lambda_1, ..., \lambda_p)$ y μ es la distribución de Haar en $O(p)$.

En particular, cuando A es invariante bajo conjugaciones ortogonales, la densidad conjunta es

$$\frac{\pi^{p^2/2}}{\Gamma_p(\frac{1}{2}p)} f(L) \prod_{i<j}^{p} (\lambda_j - \lambda_i)$$

y los vectores y los valores propios de A son independientes y los vectores propios tienen la distribución de Haar.

Recordemos que B_n **matriz** $n \times n$ **real simétrica gaussiana** GOE(n)**:**

$$B_n(j,k), 1 \leq j \leq k \leq n \quad \text{v.a. independientes}$$

$$B_n = \begin{bmatrix} B_n(1,1) & \cdots & B_n(1,n) \\ \vdots & & \vdots \\ B_n(n,1) & \cdots & B_n(n,n) \end{bmatrix}$$

$$\begin{aligned} &B_n(j,k) = B_n(k,j), \\ &B_n(j,k) \sim N(0,1), \ j \neq k, \\ &B_n(j,j) \sim N(0,2). \end{aligned}$$

Teorema: Una matriz aleatoria $n \times n$ simétrica B_n es $GOE(n)$ si y sólo si las siguientes dos condiciones se cumplen:

a) B_n es invariante bajo conjugación ortogonal (simetría):

$$\text{las distribuciones de } OB_nO^T \text{ y } B_n \text{ son las mismas} \quad \forall O \in \mathbb{O}(n)$$

b) B_n tiene entradas independientes en la diagonal y arriba de la diagonal (matriz de Wigner)

Sobre la densidad conjunta de valores propios de GOE (n)

- Fórmula de Weyl: valores propios $\lambda_{n,1} > ... > \lambda_{n,n}$ de B_n

$$f_{\lambda_{n,1},...,\lambda_{n,n}}(x_1,...,x_n) = k_n \underbrace{\left[\prod_{j=1}^{n} \exp\left(-\frac{1}{4}x_j^2\right)\right]}_{\text{independencia}} \underbrace{\left[\prod_{j<k} |x_j - x_k|\right]}_{\text{repulsión}}$$

- Si los valores propios fueran independientes:

$$f_{\lambda_{n,1},...,\lambda_{n,n}}(x_1,...,x_n) = f_{\lambda_{n,1}}(x_1) \cdots f_{\lambda_{n,n}}(x_n).$$

- Todo lo contrario $\lambda_{n,1},...,\lambda_{n,n}$ son **fuertemente dependientes**. Hay contribuciones de independencia y **repulsión**.
- Fenómeno general en matrices aleatorias con entradas continuas: Jacobiano función del determinante de Vandermont: $x = (x_1,...,x_n)$

$$\Delta(x) = \det\left(\left\{x_j^{k-1}\right\}_{j,k=1}^{n}\right) = \prod_{j<k}(x_j - x_k).$$

Sobre la densidad conjunta de valores propios de GUE (n)

- Fórmula de Weyl: valores propios $\lambda_{n,1} > ... > \lambda_{n,n}$ de B_n

  Ensambles Gaussiano Unitario $\beta = 2$

$$f_{\lambda_{n,1},...,\lambda_{n,n}}(x_1,...,x_n) = k_{n,\beta} \left[\prod_{j=1}^{n} \exp\left(-\frac{\beta}{4}x_j^2\right)\right] \left[\prod_{j<k} |x_j - x_k|^{\beta}\right]$$

- $\beta = 1, 2$ o 4.
- Referencia recomendada: Sec. 2.5 libro Anderson, *et a*l. (2010).

2.2.3. Sobre el teorema de Wigner

La motivación de Wigner

- *Wigner observó que un núcleo pesado es una gota líquida compuesta de muchas partículas con fuertes interacciones desconocidas.*
- *Modelo propuesto: valores propios de una matriz aleatoria.*
- *¿Cuál es la distribución de estas partículas?.*
- *¿Qué tipo de matrices aleatorias deben usarse?.*
- Lecturas recomendadas de Wigner sobre Matrices aleatorias:
  - *Annals of Mathematics*, 1955, 1958.
  - Random Matrices in Physics, SIAM Review, (7th Von Neumann Lecture) en 1967.
- Wigner (1902-1995) gano Premio Nobel en 1963 por sus estudios sobre el núcleo atómico.

En general el estudio asintótico de $\widehat{F}_n$ no es trivial debido a la fuerte dependencia entre los valores propios.

- **Teorema Wigner (1955):** $\forall\ f \in C_b(\mathbb{R})$ y $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(\left|\int f(x)\mathrm{d}\widehat{F}_n(x) - \int f(x)\mathrm{w}(x)\mathrm{d}x\right| > \varepsilon\right) = 0$$

  donde $\mathrm{w}(x)\mathrm{d}x$ es la *distribución del semicírculo en* $(-2, 2)$ con densidad

$$\mathrm{w}(x) = \frac{1}{2\pi}\sqrt{4 - x^2}, \quad |x| \leq 2.$$

- **En otras palabras**: la distribución empírica espectral $\widehat{F}_n$ converge (en cierto sentido) a la distribución del semicírculo en $(-2, 2)$.
- Distribución del semicírculo

$$\mathrm{w}(x) = \frac{1}{2\pi}\sqrt{4 - x^2}, \quad |x| \leq 2$$

- Es la distribución gaussiana en probabilidad libre.
- Es la distribución límite en el teorema central del límite libre.

- Da lugar al movimiento browniano libre.

$$\mathrm{w}_t(x) = \frac{1}{2\pi}\sqrt{4t - x^2}, \quad |x| \leq 2\sqrt{t}$$

- Independencia libre, independencia libre asintótica (Dan-Virgil Voiculescu, 90´s).

- Divisibilidad infinita libre y procesos de Lévy libres

2.2.4. Relaciones inesperadas entre Matrices Aleatorias y otras áreas

???Añadir lo de la función zeta de Riemann.

2.2.5. Teorema de Marchenko-Pastur

Falta añadir esto.

2.2.6. Tracy-Widom

Falta esto

Añadir biografías.

Capítulo 3

Matrices Aleatorias y comunicación inalámbrica

Mario Díaz

Índice de Figuras

Índice de Tablas

Bibliografía

[BS10] Zhidong Bai and Jack W Silverstein. *Spectral analysis of large dimensional random matrices*, volume 20. Springer, 2010.

[DPA17] Mario Diaz and Victor Pérez-Abreu. On the capacity of block multiantenna channels. *IEEE Transactions on Information Theory*, 2017.

[DR11] Armando Domínguez and Alfonso Rocha. El teorema de Wigner para matrices aleatorias. (52):31–51, 2011.

[Ede89] Alan Edelman. Eigenvalues and Condition Numbers of Random Matrices. *PhD thesis at the MIT*, 1989.

[ER60] Paul Erdös and Alfréd Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5(1):17–60, 1960.

[KS13] Michael Krivelevich and Benny Sudakov. The phase transition in random graphs: A simple proof. *Random Structures & Algorithms*, 43(2):131–138, 2013.

[Mec09] Mark W Meckes. Some results on random circulant matrices. In *High dimensional probability V: the Luminy volume*, pages 213–223. Institute of Mathematical Statistics, 2009.

[MS12] James A Mingo and Roland Speicher. Sharp bounds for sums associated to graphs of matrices. *Journal of Functional Analysis*, 5(262):2272–2288, 2012.

[Sha48] Claude E Shannon. A mathematical theory of communication, part i, part ii. *Bell Syst. Tech. J.*, 27:623–656, 1948.

[Tao13] Terence Tao. Topics in Random Matrix Theory. In *Deutsche Mathematiker-Vereinigung.* Springer-Verlag Berlin Heidelberg, 2013.

[Tel99] Emre Telatar. Capacity of multi-antenna gaussian channels. *Transactions on Emerging Telecommunications Technologies*, 10(6):585–595, 1999.

[VDH09] Remco Van Der Hofstad. Random graphs and complex networks. *Available on http://www. win. tue. nl/rhofstad/NotesRGCN. pdf*, 11, 2009.

Índice Analítico