



Centro de Investigación en Matemáticas, A.C.

**Sobre Aproximaciones Gaussianas a las
Matrices Aleatorias con Distribución de Haar**

Tesina

que para obtener el Grado de:

Maestro en Ciencias con Orientación en Probabilidad y Estadística

Presenta:

Dialid Santiago Ramírez

Director de tesina:

Dr. Víctor Manuel Pérez-Abreu Carrión

Guanajuato, Guanajuato, México

11 Julio de 2011

Integrantes del jurado

Presidente: Dr. Alfonso Rocha Arteaga

Secretario: Dr. Eduardo Dueñez Guzmán

Vocal y director de la tesina: Dr. Víctor Manuel Pérez Abreu Carrión

Asesor:

Dr. Víctor Manuel Pérez-Abreu Carrión.

Sustentante:

Dialid Santiago Ramírez.

Dedicatoria

*Con gran cariño y gratitud, a mis padres Dialid e Isaías
por su interminable apoyo y paciencia.*

*Con todo mi amor a Javier por ser mi amor,
mi esperanza y mi felicidad.*

Agradecimientos

A mis padres y a mi querida hermanita Astrid, por todo el amor y el respaldo que me ha brindado, sin ustedes nada de esto hubiera sido posible.

A Javier, por su apoyo, su compañía y su paciencia durante todos y cada uno de los días de la realización de este trabajo.

Un agradecimiento con profunda admiración a mi asesor, el Dr. Víctor Pérez-Abreu Carrión, por confiar en mí al darme la oportunidad de ser su estudiante, por sus consejos, su dedicación y su gran paciencia durante la dirección de este trabajo.

A mi querida amiga Rosy Dávalos, por alentarme a empezar la maestría, por ser mi ejemplo de trabajo y fortaleza, pero sobre todo por su amistad sincera.

A mis amigos Ehyter, Elena, José Manuel y Rosana por brindarme su amistad y por los momentos inolvidables que pasamos juntos durante estos dos años.

A mis sinodales, el Dr. Alfonso Rocha Arteaga y el Dr. Eduardo Dueñez Guzmán, por sus valiosos comentarios y por tomarse el tiempo para leer y revisar este trabajo.

A quienes fueron mis profesores, por su enseñanza y consejo, en especial a Victor Rivero y a Joaquín Ortega, por su apoyo a lo largo de estos dos años, por darme ánimo y aliento y por creer en mí.

Al Consejo Nacional de Ciencia y Tecnología y al Centro de Investigación en Matemáticas A.C. por el apoyo económico recibido para la realización de mis estudios de maestría.

Índice general

Introducción	1
1. Preliminares	5
1.1. Ensamblés Gaussianos	6
1.1.1. Ensamblés Hermitianos	6
1.1.2. Ensamblés de Ginibre	8
1.2. Resultados Asintóticos sobre el Espectro	9
1.2.1. Ley del Círculo	10
1.2.2. Ley de Tracy-Widom	11
1.3. Matrices de covarianza muestral	12
1.3.1. Ley de Marchenko-Pastur	12
1.3.2. Límite de los valores propios extremos	13
1.4. Matrices Aleatorias con Distribución de Haar	13
2. Generación de Matrices Aleatorias de Haar	17
2.1. Métodos de Generación y la Descomposición QR	18
2.1.1. Método de Gram-Schmidt	19
2.1.2. Reflexiones Householder	22
2.1.3. Observaciones sobre la Descomposición QR	24
2.2. Una Interpretación Algebraica	27
3. Aproximación de Matrices Aleatorias de Haar	31
3.1. Sobre un Problema Planteado por Diaconis	32
3.2. Solución de Jiang	33
3.2.1. Resultados Auxiliares	35
3.2.2. Pruebas de los Resultados Principales	42
3.3. Convergencia del Máximo de Diferencias de Entradas	43

3.3.1. Resultado Principal	43
3.3.2. Resultados Auxiliares	44
3.3.3. Prueba del Resultado Principal	54
4. Consecuencias sobre la Distribución Espectral Asintótica de Bloques	57
4.1. Ley de Grandes Números para el Radio Espectral	58
4.2. Ley de Tracy-Widom	59
4.3. Ley del Círculo	62
4.4. Ley de Marchenko-Pastur	63
A. Resultados sobre la Distribución Espectral de $Z'Z$	67
B. Resultados sobre Cotas y Aproximaciones para Colas	69
C. Resultados Técnicos	73

Introducción

La Teoría de Matrices Aleatorias tiene sus inicios en 1928 con los trabajos de Wishart [71] en el contexto de estadística matemática. Sin embargo, fue hacia finales de los años cincuenta y principios de los sesenta que esta área recibió un impulso importante, debido a los trabajos en física de Wigner [72, 74, 75, 76], Dyson [20, 21] y Mehta [50], entre otros. El interés en el área surgió gracias al estudio de las líneas de resonancia de los núcleos pesados, así como de algunos otros sistemas cuánticos donde el hamiltoniano es desconocido y difícil o imposible de determinar con exactitud. Este desconocimiento acerca los detalles de las interacciones de los sistemas motivó la idea de prescindir de ellos y realizar una descripción estocástica considerando conjuntos de matrices, denominados ensambles, con entradas aleatorias y que obedecen ciertas restricciones derivadas de la simetría. En 1962 Dyson [20] presentó una clasificación general de estos conjuntos de matrices de acuerdo con dichas simetrías.

Entre los objetos de estudio más importantes en matrices aleatorias están los llamados ensambles de matrices: sucesiones de matrices de dimensión creciente dotadas de una medida de probabilidad. Algunos aspectos de interés sobre los elementos de estos conjuntos son: su distribución espectral, la correlación de sus valores propios y las separaciones entre ellos, así como al estudio de otras de sus estadísticas relevantes, en particular la obtención de resultados límites para cuando la dimensión de las matrices crece a infinito.

En este trabajo estamos interesados en una clase particular de ensambles de matrices aleatorias, aquellos formados por matrices con distribución de Haar. Las matrices de estos ensambles pertenecen a los grupos compactos clásicos: $\mathcal{O}(n)$ ortogonal, $\mathcal{U}(n)$ unitario y $\mathcal{Sp}(n)$ simpléctico y su distribución de probabilidad matricial es la correspondiente medida de Haar normalizada definida sobre cada grupo.

Una de las razones por las que estudiar matrices aleatorias de Haar resulta de interés es la amplia variedad de aplicaciones donde sus eigenvalores o sus entradas desempeñan un papel relevante. Como ejemplos - y lejos de ser extensivos - pueden mencionarse aplicaciones en mecánica cuántica [50], teoría de números [45, 46], telecomunicaciones [62] y teoría de operadores [6, 26]. Algunas de estas aplicaciones involucran el uso de pruebas estadísticas [18] o aproximaciones numéricas, lo que hace patente la necesidad de contar con métodos para generar matrices aleatorias con distribución de Haar.

De igual importancia es de interés contar con resultados teóricos acerca de la distribución espectral límite para ensambles de matrices aleatorias de Haar, en forma análoga a los que han sido estudiados para otros tipos de matrices aleatorias. Entre ellos destacan tres resultados clásicos y pioneros para los ensambles de matrices hermitianas con entradas gaussianas independientes: (1) La Ley del Semicírculo, probada por Wigner [73] en 1955, que establece que la distribución espectral empírica de matrices de los Ensamblés gaussianos converge a la ley de semicírculo; (2) La Ley de Marchenko-Pastur [49], probada en 1967, que describe la convergencia de la distribución espectral empírica de matrices de covarianza, y (3) el trabajo de Tracy y Widom [67] en 1994, sobre la convergencia del máximo eigenvalor de matrices gaussianas hacia cierta distribución, que se conoce ahora como Ley de Tracy-Widom.

La relación entre matrices con entradas gaussianas e independientes y matrices con distribución de Haar es un tema de interés que tiene sus orígenes en el trabajo de Borel [7] de 1906. En ese trabajo Borel demuestra el resultado que ahora es conocido como Teorema de Poincaré, el cual tiene como consecuencia que si $\Gamma_n = (\gamma_{ij})_{1 \leq i, j \leq n}$ tiene distribución de Haar, entonces su primera entrada reescalada por \sqrt{n} converge débilmente a una ley gaussiana estándar cuando n crece a infinito. Desde entonces se han probado varios resultados que ponen de manifiesto la similitud en el comportamiento de ambos tipos de matrices cuando sus dimensiones son suficientemente grandes, a continuación. Entre los que podemos mencionar algunos de la década de los 90's: Diaconis y Shahshahani [16] y Rains [57] estudiaron esta similitud en el comportamiento de la esperanza de las trazas, mientras que Johanson [41] estudió la razón de convergencia de la distribución de las trazas. En 2002 D'Aristotile, Diaconis y Newman [11] probaron que la combinación lineal de entradas de matrices aleatorias de Haar también converge en distribución a una ley gaussiana. Estos resultados motivaron el estudio de aproximaciones más generales para matrices aleatorias con distribución de Haar vía matrices aleatorias gaussianas, así como la posible extensión de resultados que se han probado para matrices gaussianas al caso de matrices de Haar.

El objetivo principal de esta tesina es presentar una revisión detallada de artículos de investigación recientes acerca de generación y aproximación de matrices aleatorias con distribución de Haar. Estos artículos también incluyen resultados sobre la distribución espectral asintótica de bloques de matrices aleatorias con distribución de Haar cuyas demostraciones se basan en las aproximaciones.

En particular, se presenta una exposición detallada de resultados recientes acerca de dos enfoques de aproximación general de matrices aleatorias de Haar por medio de matrices con entradas gaussianas independientes. El primero consiste en estudiar el comportamiento de la distancia variacional entre las distribuciones matriciales correspondientes, cuando la dimensión crece. Por otra parte, el segundo se basa en el comportamiento asintótico de la máxima diferencia entre las entradas de ambas matrices.

La estructura de esta tesina es la siguiente: en el Capítulo 1 se hace una breve revisión de algunos conceptos elementales y resultados clásicos sobre los llamados ensambles gaussianos: Ensamble

Gaussiano Ortogonal (GOE), Ensamble Gaussiano Unitario (GUE) y Ensamble Gaussiano Simpléctico (GSE). Los elementos de estos ensambles son matrices hermitianas con entradas independientes (reales, complejas, o cuaternios, respectivamente) y con distribución gaussiana, cuya distribución conjunta posee la propiedad de invarianza bajo la acción por conjugación de elementos del grupo de matrices ortogonales, unitarias y simplécticas, respectivamente. Así mismo, presentamos definiciones y resultados que serán de utilidad a lo largo del trabajo.

En el Capítulo 2 se exponen algunos métodos para generar matrices aleatorias con distribución de Haar sobre los tres grupos compactos clásicos. La idea fundamental detrás de estos métodos es obtener la descomposición QR de una matriz con entradas independientes con ley gaussiana y observar que Q tiene distribución de Haar. Existen diferentes métodos para obtener dicha factorización, pero en este trabajo nos concentramos en dos de ellos: el Proceso de Ortonormalización de Gram-Schmidt y el Método de Reflexiones Householder. El primero debido a que será de especial importancia en una de las aproximaciones y el segundo por que debido a su eficiencia computacional es el más utilizado en la práctica.

El Capítulo 3 se inicia exponiendo un problema planteado por Diaconis [17] acerca de encontrar el máximo orden para las dimensiones del bloque superior izquierdo de una matriz aleatoria de Haar de tal manera que su distribución conjunta se aproxime en algún sentido a la distribución conjunta de una matriz aleatoria, del mismo tamaño, con entradas i.i.d. con distribución gaussiana estándar, cuando la dimensión de las matrices tiende a infinito. Después se enuncian dos Teoremas obtenidos por Jiang [38] en 2006, que da una respuesta al problema antes mencionado. Las demostraciones de ambos resultados se exponen clara y detalladamente en este capítulo. Posteriormente se estudia otro enfoque para estudiar la “cercanía” que existe entre bloques de matrices aleatorias de Haar y matrices con entradas i.i.d. con ley gaussiana del tamaño del bloque, cuando la dimensión es suficientemente grande. La idea consiste en estudiar el comportamiento límite del máximo de la norma de las diferencias entre sus entradas. En 2006 Jiang [37] prueba que en el caso ortogonal dicho máximo converge a cero en probabilidad, bajo ciertas condiciones sobre el orden del tamaño del bloque. En 2010 Jiang [40] generalizó su resultado a los casos unitario y simpléctico. El objetivo central de la última sección del Capítulo 4 es presentar esta generalización así como su demostración.

En el Capítulo 4 se estudian algunas implicaciones relevantes de las aproximaciones antes mencionadas. Específicamente, estas implicaciones se refieren a la distribución límite espectral de bloques de matrices de Haar.

Con el objeto de hacer la lectura de este trabajo fluida y a la vez auto contenida, se presentan tres apéndices, donde se exponen resultados que se usan a lo largo del trabajo, pero que pueden ser estudiados de manera independiente de los resultados principales. El primero contiene resultados acerca de la distribución espectral de matrices del tipo $Z'Z$ donde Z es una matriz con entradas independientes

e idénticamente distribuidas, el segundo se refiere a resultados que dan cotas o aproximaciones para la cola de variables aleatorias con distribución gaussiana estándar y finalmente el tercero contiene resultados de propiedades varias de caracter técnico usadas en el trabajo.

Capítulo 1

Preliminares

A lo largo de este trabajo se considerarán las siguientes convenciones y notación:

Conjuntos y números

- \mathbb{R}, \mathbb{C} denota el conjunto de números reales y complejos respectivamente.
- \mathbb{H} denota al conjunto de cuaterniones, también llamados cuaternios, es decir

$$\mathbb{H} = \{x_1 + ix_2 + jx_3 + kx_4 : x_i \in \mathbb{R} \ i = 1, 2, 3, 4\}$$

donde i, j, k satisfacen que $i^2 = j^2 = k^2 = ijk = -1$.

- \mathbb{F}_β denotará a \mathbb{R}, \mathbb{C} o \mathbb{H} , para $\beta = 1, 2$ y 4 respectivamente.

Grupos

- Los grupos que serán de interés en este trabajo, así como la notación que se seguirá de aquí en adelante, se recuerdan en la siguiente definición.

Definición 1. Los grupos compactos clásicos de matrices son:

- El grupo Ortogonal de dimensión n , es el grupo formado por todas las matrices reales ortogonales y de dimensión $n \times n$, es decir

$$\mathcal{O}(n) := \{\mathbf{O} \in \mathbb{R}^{n \times n} \mid \mathbf{O}^t \mathbf{O} = \mathbf{I}\}.$$

- El grupo Unitario de dimensión n , es el grupo formado por todas las matrices complejas unitarias y de dimensión $n \times n$, es decir

$$\mathcal{U}(n) := \{\mathbf{U} \in \mathbb{C}^{n \times n} \mid \mathbf{U}^* \mathbf{U} = \mathbf{I}\}.$$

- El grupo Simpléctico de dimensión n , es el grupo formado por todas las matrices cuaterniónicas unitarias y de dimensión $n \times n$, es decir

$$Sp(n) := \{\mathbf{S} \in \mathbb{H}^{n \times n} | \mathbf{S}^* \mathbf{S} = \mathbf{I}\}.$$

- Cuando se quiera hacer referencia a cualquiera de estos grupos (indistintamente) se usará la notación $\mathcal{G}_\beta(n)$, esto es

$$\mathcal{G}_\beta(n) := \{\mathbf{G} \in \mathbb{F}_\beta^{n \times n} | \mathbf{G}^* \mathbf{G} = \mathbf{I}\},$$

donde β puede ser 1,2, o 4.

Probabilidad

- $X \stackrel{\mathcal{L}}{=} Y$ se usará para denotar igualdad en distribución.
- $N(\mu, \sigma)$ denota una distribución gaussiana de media μ y varianza σ^2 .
- $\text{NF}_\beta(0, 1)$ denota a la distribución gaussiana estándar sobre el campo \mathbb{F}_β , esto es si $Z \sim \text{NF}_\beta(0, 1)$, entonces

$$Z := \begin{cases} \xi_1, & \text{si } \beta = 1; \\ \xi_1 + i\xi_2/\sqrt{2}, & \text{si } \beta = 2; \\ \xi_1 + i\xi_2 + j\xi_3 + k\xi_4/2, & \text{si } \beta = 4. \end{cases}$$

donde $\{\xi_i\}_{i=1}^4$ son variables aleatorias independientes con distribución normal estándar.

1.1. Ensamblés Gaussianos

1.1.1. Ensamblés Hermitianos

Estos ensambles están formados por matrices Hermitianas (o Hermíticas) cuyas entradas son variables aleatorias en \mathbb{F}_β , independientes y con distribución gaussiana $\text{NF}_\beta(0, 1)$ para $\beta = 1, 2$ y 4, respectivamente. Las matrices de estos ensambles se caracterizan por el hecho de que su distribución conjunta matricial resulta invariante bajo multiplicación por matrices unitarias apropiadas.

Definición 2. Una matriz aleatoria simétrica $\mathbf{Z}_n = (z_{ij})_{1 \leq i, j \leq n}$ pertenece al Ensamble Gaussiano Ortogonal, al que denotaremos por $\text{GOE}(n)$ si sus entradas son independientes y con distribución

$$\frac{1}{\sqrt{\pi}} \exp\{-z_{ij}^2\} \quad y \quad \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{z_{jj}^2}{2}\right\},$$

cuando $i = j$ y cuando $i \neq j$ respectivamente.

Una forma para construir matrices del $\text{GOE}(n)$ es considerar matrices \mathbf{Y} de tamaño $n \times n$ de entradas independientes y con distribución gaussiana $\mathcal{NF}_1(0, 1)$ y definir

$$\mathbf{Z}_n = \frac{1}{2}(\mathbf{Y} + \mathbf{Y}^t),$$

así $\mathbf{Z} \in \text{GOE}(n)$.

De la definición anterior resulta fácil calcular explícitamente la función de densidad de probabilidad conjunta de \mathbf{Z}_n con respecto a la medida de Lebesgue es

$$\begin{aligned} P(Z) &= \prod_{i=1}^n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z_{ii}^2}{2}\right) \prod_{i < j \leq n} \frac{1}{\sqrt{\pi}} \exp\{-z_{ij}^2\} \\ &= A_n^1 \prod_{i,j=1}^n \exp\left\{-\frac{z_{ij}^2}{2}\right\} \\ &= A_n^1 \exp\left\{-\frac{1}{2} \text{tr}(Z^2)\right\}, \end{aligned}$$

donde A_n^1 denota una constante de normalización apropiada y $\text{tr}(\cdot)$ al operador traza.

Definición 3. Una matriz aleatoria Hermitiana $\mathbf{Z}_n = (z_{ij})_{1 \leq i,j \leq n}$ pertenece al Ensamble Gaussiano Unitario, al que denotaremos por $\text{GUE}(n)$ si sus entradas $z_{ij} = x_{ij} + iy_{ij}$ son independientes y con distribución

$$N(0, 1/\sqrt{2}) \quad \text{y} \quad N(0, 1/2) + iN(0, 1/2),$$

para $i = j$ (note que estos elementos son reales) y cuando $i \neq j$ respectivamente.

Tal como en el caso del $\text{GOE}(n)$, una forma de obtener matrices del $\text{GUE}(n)$ es considerar matrices \mathbf{Y} de tamaño $n \times n$ con entradas i.i.d. y con distribución $\mathcal{NF}_2(0, 1) = N(0, 1/\sqrt{2}) + iN(0, 1/\sqrt{2})$ y luego definir

$$\mathbf{Z}_n = \frac{1}{2}(\mathbf{Y} + \mathbf{Y}^*).$$

En este caso la función de densidad de probabilidad conjunta de \mathbf{Z}_n será

$$\begin{aligned} P(Z) &= \prod_{i=1}^n \frac{1}{\sqrt{\pi}} \exp(-x_{ii}^2) \prod_{1 \leq i < j \leq n} \frac{2}{\pi} \exp\{-2|z_{ij}|^2\} \\ &= A_n^2 \prod_{i,j=1}^n \exp\{-|z_{ij}|^2\} \\ &= A_n^2 \exp\{-\text{tr}(Z^2)\}, \end{aligned}$$

donde A_n^2 denota una constante de normalización adecuada.

Definición 4. Una matriz Hermitiana \mathbf{Z}_n de tamaño $n \times n$ pertenece al Ensamble Gaussiano Simplético, que se denotará por $\text{GSE}(n)$, si sus entradas son variables aleatorias cuaterniónicas independientes cuyas entradas de la diagonal tienen distribución $N(0, 1/2)$, mientras que aquellas fuera de la diagonal son de la forma $z_{ij} + jw_{ij}$, donde $z_{ij} = u_{ij} + iv_{ij}$ y $w_{ij} = u'_{ij} + iv'_{ij}$ tienen distribución $N(0, 1/2\sqrt{2}) + iN(0, 1/2\sqrt{2})$.

Tal como en los casos anteriores, es fácil ver que

$$P(\mathbf{Z}) = A_n^4 \exp \{-2 \text{tr}(\mathbf{Z}^2)\},$$

donde A_n^4 denota una constante de normalización.

1.1.2. Ensamblados de Ginibre

A mediados de la década de los 60's Jean Ginibre [28] introdujo tres ensambles de matrices aleatorias, que ahora se denominan "Ensamblados de Ginibre" y se denotan por $\text{GinOE}(n)$, $\text{GinUE}(n)$ y $\text{GinSE}(n)$, respectivamente. Los elementos de estos ensambles se caracterizan por tener entradas independientes y con distribución gaussiana estándar $\mathbb{N}\mathbb{F}_\beta$, para $\beta = 1, 2$ y 4 respectivamente.

Sin pérdida de generalidad, considere el caso unitario. Sea $\mathbf{Z} \in \text{GinUE}(n)$, su distribución conjunta esta dada por

$$\begin{aligned} P(\mathbf{Z}) &= \frac{1}{\pi^{n^2}} \prod_{i,j=1}^n \exp \{-|z_{ij}|^2\} = \frac{1}{\pi^{n^2}} \exp \left\{ - \sum_{i,j=1}^n |z_{ij}|^2 \right\} \\ &= \frac{1}{\pi^{n^2}} \exp \{-\text{tr}(\mathbf{Z}^* \mathbf{Z})\}. \end{aligned}$$

Lema 1.1. *La distribución conjunta de elementos de $\text{GinUE}(n)$ es invariante bajo multiplicación, izquierda y derecha de elementos del grupo $\mathcal{U}(n)$, i.e. para toda $\mathbf{Z} \in \text{GinUE}(n)$ se cumple que*

$$\mathbf{UZ} \stackrel{\mathcal{L}}{=} \mathbf{ZV} \stackrel{\mathcal{L}}{=} \mathbf{Z}, \tag{1.1}$$

para cualesquiera $\mathbf{U}, \mathbf{V} \in \mathcal{U}(n)$.

Demostración. Para probar la invarianza por la izquierda hay que mostrar que

$$P(\mathbf{Z}) = P(\mathbf{UZ}), \text{ para toda } \mathbf{U} \in \mathcal{U}(n),$$

y que el Jacobiano del mapeo

$$\mathbf{Z} \mapsto \mathbf{UZ}, \tag{1.2}$$

visto como una transformación en \mathbb{C}^{n^2} , es igual a uno.

En efecto, se tiene que

$$P(\mathbf{UZ}) = \frac{1}{\pi^{n^2}} \exp\{-\operatorname{tr} \mathbf{Z}^* \mathbf{U}^* \mathbf{UZ}\} = \frac{1}{\pi^{n^2}} \exp\{-\operatorname{tr} \mathbf{Z}^* \mathbf{Z}\} = P(\mathbf{Z})$$

con lo que (1.1) queda probado.

Por otra parte, el mapeo (1.2) es isomorfo a

$$\mathbf{X} = \bigoplus_1^n \mathbf{U}.$$

con lo que \mathbf{X} es una matriz unitaria de tamaño $n^2 \times n^2$ y por tanto su determinante es igual a 1 y se obtiene el resultado que se quería.

Análogamente se prueba el caso la invarianza bajo multiplicación por la derecha y se tiene el resultado. ■

1.2. Resultados Asintóticos sobre el Espectro

En esta sección se enuncian algunos resultados clásicos en teoría de matrices aleatorias.

Definición 5. Sea \mathbf{Z}_n una matriz aleatoria de tamaño $n \times n$ con eigenvalores λ_i . Se dice que $F_{\mathbf{Z}_n}$ es la distribución espectral empírica de \mathbf{Z}_n si es aquella medida de probabilidad aleatoria que asigna igual masa a cada uno de sus eigenvalores.

Note que si $\lambda_1, \dots, \lambda_n$ son los n eigenvalores de \mathbf{Z}_n , la definición anterior corresponde a la distribución

$$F_{\mathbf{Z}_n}(x, y) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}\{\Re(\lambda_i) \leq x, \Im(\lambda_i) \leq y\}.$$

para cualesquiera $x, y \in \mathbb{R}$.

Por otra parte, en el caso de que todos los λ_i son reales (e.g. si Z pertenece al GOE), la función de distribución espectral empírica corresponderá a la función de distribución (unidimensional)

$$F_{\mathbf{Z}_n}(x) = \frac{1}{n} \sum_{i=1}^m \mathbf{1}\{\lambda_i \leq x\},$$

para todo $x \in \mathbb{R}$.

Es de interés estudiar la convergencia de la sucesión de distribuciones espectrales empíricas $\{F_{\mathbf{Z}_n}\}$ correspondientes a sucesiones de matrices aleatorias $\{\mathbf{Z}_n\}$ cuya dimensión crece a infinito.

La distribución límite

$$F = \lim_{n \rightarrow \infty} F_{\mathbf{Z}_n},$$

que puede ser degenerada en el sentido de que su masa total sea menor que 1 (cuando algunos de los eigenvalores tienden a $\pm\infty$), se denomina *Distribución espectral límite* de la sucesión $\{\mathbf{Z}_n\}$.

1.2.1. Ley del Círculo

Este resultado se refiere a la convergencia de la distribución espectral de matrices aleatorias normalizadas hacia la distribución uniforme sobre el disco unitario en el plano complejo.

Formalmente, sea $\mathbf{A}_n = (a_{ij})$ una matriz aleatoria de tamaño $n \times n$ con entradas i.i.d. de esperanza cero y varianza $\sigma^2 < \infty$ y considere

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n,$$

los eigenvalores de la matriz $\mathbf{Z}_n = \frac{1}{\sigma\sqrt{n}}\mathbf{A}_n$. Desde 1950 se conjeturó que la distribución espectral empírica de \mathbf{Z}_n , es decir

$$F_{\mathbf{Z}_n}(s, t) := \frac{1}{n} \# \{k \leq n \mid \Re(\lambda_k) \leq s, \Im(\lambda_k) \leq t\} \quad (1.3)$$

converge a la distribución uniforme sobre el círculo unitario.

Entre los primeros trabajos donde se realizaron progresos para probar esta conjetura podemos mencionar el trabajo de Mehta [50] en 1960, así como el trabajo inédito de Silverstein citado en 1986 por Hwang [35]. En ambos se trató el caso en el que las entradas de la matriz son variables aleatorias complejas i.i.d. con distribución normal estándar y la demostración se basa en utilizar la expresión explícita de la densidad conjunta de los eigenvalores complejos de la matriz aleatoria, es decir

$$c \prod_{j \neq k} |\lambda_j - \lambda_k|^2 \exp \left\{ -\frac{1}{2} \sum_{k=1}^n |\lambda_k|^2 \right\},$$

que fue calculada en 1965 por Ginibre [28].

El primer intento por probar que la conjetura es válida bajo condiciones más generales fue hecho en 1984 por Girko [29, 30]. En ese trabajo se muestra que el resultado es válido bajo las condiciones de que las densidades tanto de la parte real como imaginaria de las entradas fuesen acotadas y con momentos finitos hasta cierto orden. Sin embargo, sus pruebas presentan algunos pasos poco claros e incluso han sido cuestionadas. De manera que el problema siguió considerándose abierto hasta 1997, cuando Bai [2], siguiendo las ideas de Girko, presenta la primera demostración rigurosa bajo estas condiciones. También en 1997 Edelman [24] demostró la convergencia del valor esperado de la función de distribución empírica.

En trabajos más recientes como el de Pan y Zhou [54] en 2007 y Tao y Vu [66] en 2008 se ha eliminado la condición sobre las densidades y se ha debilitado la condición sobre la existencia de momentos.

1.2.2. Ley de Tracy-Widom

Este resultado se refiere a la distribución límite de los eigenvalores extremos de matrices aleatorias y para enunciarlo es necesario introducir a la función $q(x)$ definida como la solución de la siguiente ecuación diferencial con condición de frontera, conocida como ecuación de Painléve II,

$$q''(x) = xq(x) + 2q(x)^3 \tag{1.4}$$

$$q(x) = Ai(x), \quad \text{cuando } x \rightarrow +\infty,$$

donde $Ai(x)$ es la función de Airy.

El primer trabajo al respecto data de 1994 y se debe a Tracy y Widom [67, 68, 69], quienes establecieron el siguiente Teorema sobre la convergencia en distribución para el máximo eigenvalor de matrices aleatorias en los ensambles gaussianos, Ortogonal, Unitario y Simplético hacia ciertas leyes límite, que ahora son llamadas distribuciones de Tracy Widom.

Teorema 1.2. (Tracy-Widom 1994) Sea Z_n una matriz aleatoria del ensamble gaussiano Hermitiano GOE, (GUE, y GSE respectivamente) y $\lambda_{\text{máx}}$ su eigenvalor máximo entonces

$$n^{2/3}(\lambda_{\text{máx}} - 2)$$

converge en distribución a la función F_β con $\beta = 1, 2, 4$ respectivamente, donde

$$\begin{aligned} F_1(s) &= \exp \left\{ -\frac{1}{2} \int_s^\infty q(x) + (x-s)q(x)^2 dx \right\}, \\ F_2(s) &= \exp \left\{ -\int_s^\infty (x-s)q(x)^2 dx \right\}, \\ F_4\left(\frac{s}{\sqrt{2}}\right) &= \cosh \left(-\frac{1}{2} \int_s^\infty q(x) dx \right) F_2(s)^{1/2} \end{aligned} \tag{1.5}$$

para $s \in \mathbb{R}$ y $q(s)$ queda definida como la solución de la ecuación (1.4).

En 2001 Johnstone [42] probó que el máximo eigenvalor de matrices de Wishart complejas converge también a una de las distribuciones de Tracy Widom y más tarde demostró el resultado análogo para el caso real. Ambos resultados se resumen en el siguiente teorema.

Teorema 1.3. (Johnstone) Sea \mathbf{W}_n una matriz de Wishart real (compleja), defina

$$\begin{aligned}\mu_{n,p} &= (\sqrt{n-1} + \sqrt{p})^2 \\ \sigma_{n,p} &= (\sqrt{n-1} + \sqrt{p}) \left(\frac{1}{\sqrt{n-1}} + \frac{1}{\sqrt{p}} \right)^{1/3},\end{aligned}$$

entonces

$$\frac{\lambda_{\text{máx}} - \mu_{n,p}}{\sigma_{n,p}}$$

converge en distribución a la ley de Tracy-Widom F_β con $\beta = 1$ para el caso real y $\beta = 2$ para el caso complejo.

1.3. Matrices de covarianza muestral

Recordemos la definición de una matriz de covarianza muestral.

Definición 6. Considere $\{x_{jk}, j, k = 1, 2, \dots\}$ un doble arreglo de variables aleatorias complejas independientes e idénticamente distribuidas con media cero y varianza σ^2 . Usando la siguiente notación

$$\mathbf{x}_j = (x_{1j}, \dots, x_{pj})^t,$$

$$\mathbf{X} = (\mathbf{x}_1, \dots, \mathbf{x}_n),$$

la matriz de covarianzas se define como

$$\mathbf{S} = \frac{1}{n-1} \sum_{k=1}^n (\mathbf{x}_k - \bar{\mathbf{x}})(\mathbf{x}_k - \bar{\mathbf{x}})^*,$$

donde $\bar{\mathbf{x}} = \frac{1}{n} \sum_{j=1}^n \mathbf{x}_j$.

La definición de matriz de covarianzas a veces se presenta como

$$\mathbf{S} = \frac{1}{n} \sum_{k=1}^n \mathbf{x}_k \mathbf{x}_k^* = \frac{1}{n} \mathbf{X} \mathbf{X}^*,$$

que es equivalente para efectos de su distribución espectral empírica con la anterior.

1.3.1. Ley de Marchenko-Pastur

Esta sección se refiere al comportamiento asintótico de la distribución espectral de los eigenvalores de matrices de covarianza muestral cuando la dimensión crece a infinito. Marchenko y Pastur [49] fueron los primeros en establecer su convergencia, hacia la distribución que más tarde fue nombrada en su honor, para un caso particular de matrices de covarianza.

Teorema 1.4. (Marchenko-Pastur) Sean $\{x_{ij}\}$ variables aleatorias complejas, independientes e idénticamente distribuidas con media cero y varianza σ^2 . Si $\frac{p}{n} \rightarrow c \in (0, \infty)$, entonces la distribución espectral empírica de \mathbf{S} tiende a la ley de Marchenko-Pastur determinada por la distribución

$$\mu(dx) = \frac{1}{2\pi cx} \sqrt{(b-x)(x-a)} \mathbf{1}_{[a,b]}(x) dx + \min\{0, 1 - 1/c\} \delta_0(dx), \quad (1.6)$$

donde

$$\begin{aligned} a &= \sigma^2(1 - \sqrt{c})^2 \\ b &= \sigma^2(1 + \sqrt{c})^2 \end{aligned}$$

1.3.2. Límite de los valores propios extremos

Teorema 1.5. (Jonsson) Sean $\{x_{ij}\}$ variables aleatorias complejas, independientes e idénticamente distribuidas con media cero y varianza σ^2 . Suponga que $\frac{p}{n} \rightarrow y \in (0, \infty)$ y que las x_{ij} tienen cuarto momento finito, entonces

$$-2y\sigma^2 \leq \liminf_{n \rightarrow \infty} \lambda_{\min}(\mathbf{S} - \sigma^2(1+y)\mathbf{I}) \leq \liminf_{n \rightarrow \infty} \lambda_{\max}(\mathbf{S} - \sigma^2(1+y)\mathbf{I}) \leq 2y\sigma^2$$

Si definimos $\lambda_{\min}(\mathbf{S})$ como el $(p - n + 1)$ eigenvalor más pequeño de \mathbf{S} cuando $p > n$, del Teorema anterior se sigue inmediatamente el siguiente resultado.

Teorema 1.6. Bajo las hipótesis del Teorema 1.5 se tienen las siguientes convergencias casi seguras,

$$\lim_{n \rightarrow \infty} \lambda_{\min}(\mathbf{S}) = \sigma^2(1 - \sqrt{y})^2 \quad (1.7)$$

y

$$\lim_{n \rightarrow \infty} \lambda_{\max}(\mathbf{S}) = \sigma^2(1 + \sqrt{y})^2 \quad (1.8)$$

1.4. Matrices Aleatorias con Distribución de Haar

La medida de Haar fue introducida en 1932 por Alfréd Haar [31] y puede interpretarse como una medida que permite asignar un “volumen invariante” a los subconjuntos de grupos topológicos localmente compactos.

Definición 7. Una medida μ sobre un grupo topológico localmente compacto \mathcal{X} es llamada medida de Haar si satisface las siguientes condiciones:

- $\mu(U) > 0$ para todo B no vacío y abierto en $\mathcal{B}(\mathcal{X})$.
- $\mu(xB) = \mu(B)$ para todo $B \in \mathcal{B}(\mathcal{X})$ y $x \in \mathcal{X}$.

Note que la medida de Haar puede interpretarse como un análogo de la medida de Lebesgue sobre \mathbb{R}^n a cualquier grupo localmente compacto \mathcal{X} , en el sentido de que esta definida sobre los Borelianos de \mathcal{X} , es regular e invariante bajo transformaciones (por la derecha ó por la izquierda, lo que en general será distinto a menos que el grupo sea conmutativo).

El siguiente teorema, cuya prueba puede verse en [34], garantiza la existencia y unicidad (salvo multiplicación por escalares) de una medida de Haar sobre cualquier grupo topológico compacto \mathcal{X} .

Teorema 1.7. *Sea \mathcal{X} un grupo topológico compacto, existe una medida de Haar sobre \mathcal{X} . Más aún, si μ y ν son medidas de Haar sobre \mathcal{X} , entonces existe una constante positiva c tal que $\mu(B) = c\nu(B)$ para todo $B \in \mathcal{B}(\mathcal{X})$.*

El resultado anterior garantiza la existencia y unicidad de la medida de Haar normalizada sobre los tres grupos compactos clásicos \mathcal{G}_β ($\beta = 1, 2, 4$), lo que permite dar la siguiente definición.

Definición 8. *Sea $\mu_{\mathcal{G}_\beta(n)}$ la medida de Haar normalizada en $\mathcal{G}_\beta(n)$ para $\beta = 1, 2, 4$. Se dice que una matriz aleatoria $\Gamma \in \mathcal{G}_\beta(n)$ tiene distribución de Haar sobre $\mathcal{G}_\beta(n)$ si su distribución matricial coincide con $\mu_{\mathcal{G}_\beta(n)}$, esto es*

$$P(\Gamma \in B) := \mu_{\mathcal{G}_\beta(n)}(B)$$

para todo $B \in \mathcal{B}(\mathcal{G}_\beta(n))$.

A continuación se enuncian algunas propiedades relevantes de las matrices aleatorias de Haar en $\mathcal{O}(n)$.

Proposición 1.8. *Sea Γ una matriz aleatoria de Haar en $\mathcal{O}(n)$, para cualquier vector \mathbf{x} de dimensión n se cumple que $\mathbf{y} = \Gamma\mathbf{x}$ sigue una distribución uniforme sobre la esfera n -dimensional S^n .*

Demostración. Sea $\mathbf{U} \in \mathcal{O}(n)$, entonces

$$\mathbf{U}\mathbf{y} = \mathbf{U}\Gamma\mathbf{x} \stackrel{\mathcal{L}}{=} \Gamma\mathbf{x} = \mathbf{y},$$

lo cual prueba que la distribución de \mathbf{y} es invariante bajo transformaciones unitarias. Usando el hecho de que esta propiedad caracteriza a la distribución uniforme sobre la esfera n dimensional [60], con lo que se sigue el resultado. ■

Proposición 1.9. *Si Γ es una matriz aleatoria de Haar en $\mathcal{O}(n)$, entonces Γ^* también lo es.*

Demostración. Sea $U \in \mathcal{O}(n)$, B en los Borelianos de $\mathcal{O}(n)$ y considere B^* el conjunto de todos los adjuntos de los elementos de B . Entonces

$$P(U\Gamma^* \in B) = P(\Gamma \in B^*U) = \mu_n(B^*) = P(\Gamma \in B^*) = P(\Gamma^* \in B),$$

lo cual implica que Γ^* tiene distribución de Haar.

Proposición 1.10. *Si Z es una matriz aleatoria de tamaño $n \times n$ con entradas i.i.d. con distribución gaussiana estándar en \mathbb{F}_β con $\beta = 1$ (2,4, respectivamente), entonces las matrices*

$$W = Z(Z^*Z)^{-1/2} \quad \text{y} \quad V = (ZZ^*)^{-1/2}Z$$

tienen distribución de Haar en $\mathcal{O}(n)$ ($\mathcal{U}(n)$, $\mathcal{S}p(n)$, respectivamente).

Demostración. En efecto, para W tenemos que para cualquier $U \in \mathcal{O}(n)$ se cumple que

$$UW = UZ(Z^*Z)^{-1/2} = UZ((UZ^*)UZ)^{-1/2} \stackrel{\mathcal{L}}{=} Z(Z^*Z)^{-1/2} = U$$

de lo cual se sigue que W tiene distribución de Haar.

Análogamente se trata el caso de V .

■

Note que la última propiedad nos proporciona un método para generar matrices aleatorias con distribución de Haar, otros métodos sobre este tema se discutirán en el Capítulo siguiente.

Capítulo 2

Generación de Matrices Aleatorias de Haar

Las matrices aleatorias con distribución de Haar en los grupos compactos clásicos han sido empleadas como herramienta de modelación en aplicaciones de diversas áreas, lo cual aumentó el interés por encontrar métodos eficientes para su generación. Entre las primeras aportaciones de este tipo puede mencionarse los trabajos de Wedderburn [70], Heiberger [32] y Stewart [64]. El primero, que trata el tema de simulación multivariada utilizando matrices aleatorias ortogonales, fue publicado en 1975. Heiberger, también motivado por el problema de simular datos de distribuciones multivariadas, propone un algoritmo (que más tarde fue corregido por Tanner y Thisted [65]) para generar matrices aleatorias ortogonales. Por su parte, en 1980 y debido a su interés en calcular estimadores condicionales, Stewart analiza la eficiencia numérica de utilizar el método de transformaciones Householder para obtener la factorización QR como un método para generar matrices aleatorias de Haar sobre $\mathcal{O}(n)$.

Entre las aplicaciones más recientes podemos mencionar el uso del polinomio característico de matrices aleatorias unitarias para modelar funciones L, que ha permitido mostrar evidencia a favor importantes conjeturas en teoría de números, así como probar resultados de interés. Los ejemplos incluyen resultados sobre el comportamiento estadístico de los ceros [52, 53, 58, 44], su distribución [45, 46], y momentos [10]. También es importante mencionar que a pesar de que existen aspectos sobre funciones L para los que la teoría de matrices aleatorias no ha dado una respuesta contundente, en muchos de ellos puede aprovecharse la generación de matrices aleatorias como un método para obtener evidencia numérica que puede ser útil para formular posibles respuestas. Esto mismo sucede en otras áreas de aplicación.

El propósito de este capítulo es presentar algunos métodos para generar matrices aleatorias con distribución de Haar sobre los grupos compactos clásicos \mathcal{G}_β $\beta = 1, 2$ y 4 , discutiendo en detalle sus ideas principales.

Adicionalmente al estudio de los métodos, en la última sección del capítulo se da una breve interpretación algebraica para ellos; con la observación de que todos ellos son casos particulares de un método más general para generar elementos de grupos compactos. Tal método se conoce como Algoritmo del Subgrupo y fue propuesto en 1987 por Diaconis y Shahshahani [14].

2.1. Métodos de Generación y la Descomposición QR

Sin pérdida de generalidad se tratará el caso unitario. La pregunta es ¿cómo generar una matriz con distribución de Haar sobre $\mathcal{U}(n)$? Primero hay que recordar que una matriz $U = (u_{ij})$ de tamaño $n \times n$ se dice unitaria si satisface la relación $U^*U = UU^* = I$, es decir

$$\sum_{k=1}^n u_{ik}^* u_{kj} = \sum_{k=1}^n \bar{u}_{ki} u_{kj} = \delta_{ij}$$

$$\sum_{k=1}^n u_{ik} u_{kj}^* = \sum_{k=1}^n u_{ik} \bar{u}_{jk} = \delta_{ij},$$

donde $U^* = (u_{ij}^*)$ denota la matriz adjunta de U , es decir $u_{ij}^* = \bar{u}_{ji}$.

Es claro que las columnas de U forman una base ortonormal de \mathbb{C}^n . Un método para construir una base de este tipo es partir de una matriz de vectores que forman una base arbitraria de \mathbb{C}^n y encontrar su factorización QR.

Definición 9. Sea A una matriz arbitraria, se dice que A tiene una descomposición o factorización QR si es posible escribirla como

$$A = QR,$$

donde Q es una matriz unitaria (ortogonal en su caso) y R es una matriz triangular superior.

A continuación se exponen algunos métodos para generar matrices aleatorias con distribución de Haar. La idea principal de estos métodos es encontrar la factorización QR de una matriz aleatoria con entradas i.i.d. con ley gaussiana estándar y surge de las siguientes observaciones:

- Sea Z una matriz aleatoria del $GinUE$, es claro que Z es de rango completo casi seguramente. Por lo que c.s. es posible encontrar una única factorización de tipo QR

$$Z = QR$$

tal que los elementos de la diagonal de R sean positivos estrictamente.

- Por otra parte, para cualquier matriz unitaria determinística H se cumple que

$$HZ \stackrel{\mathcal{L}}{=} Z,$$

debido a la invarianza bajo transformaciones unitarias de la distribución de Z .

- De las observaciones anteriores se sigue que si H es una matriz unitaria entonces,

$$QR = Z \stackrel{\mathcal{L}}{=} HZ = H(QR) = (HQ)R.$$

Los extremos de la última cadena de igualdades sugieren que $Q \stackrel{\mathcal{L}}{=} HQ$, es decir que la distribución de Q es invariante bajo transformaciones unitarias. Esta idea se enuncia de manera formal en el siguiente Teorema, cuya prueba puede verse en [22].

Teorema 2.1. *Sea $Z \in GinUE(n)$ y considere la única descomposición QR de Z tal que los elementos de la diagonal de R son estrictamente positivos, entonces Q tiene distribución de Haar en $\mathcal{U}(n)$.*

El uso de la factorización QR con propósitos numéricos está ampliamente extendido, por lo que existen varios métodos para calcularla. Entre los que podemos mencionar el proceso de ortonormalización de Gram-Schmidt, el método de transformaciones Householder, o el método de rotaciones de Givens. A continuación se describen con mayor detalle algunos de ellos.

2.1.1. Método de Gram-Schmidt

Este método consiste en generar una matriz aleatoria del ensamble $GinUE(n)$ y aplicar el proceso de ortonormalización de Gram-Schmidt a sus vectores columna. La matriz formada por los vectores resultantes será una matriz unitaria con distribución de Haar sobre $\mathcal{U}(n)$.

Método

- Sea $Z = (z_{ij})_{1 \leq i, j \leq n}$ una matriz aleatoria en $GinUE(n)$, es decir las entradas de Z son variables aleatorias i.i.d. con distribución gaussiana estándar $\mathcal{N}\mathcal{F}_2(0, 1)$. Reescribimos Z por columnas como

$$Z = (\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n),$$

esto es $\{\mathbf{z}_j\}_{1 \leq j \leq n}$ son los vectores columna de Z .

- Se definen las siguientes cantidades:

$$\mathbf{w}_1 = \mathbf{z}_1, \quad \mathbf{u}_1 = \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} \quad (2.1)$$

$$\mathbf{w}_j = \mathbf{z}_j - \sum_{k=1}^{j-1} \mathbf{u}_k \cdot (\mathbf{u}_k^* \mathbf{z}_j), \quad \mathbf{u}_j = \frac{\mathbf{w}_j}{\|\mathbf{w}_j\|} \quad j = 2, \dots, n.$$

- Se define la matriz

$$\mathbf{U} = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n). \quad (2.2)$$

Entonces afirmamos que la matriz \mathbf{U} es unitaria y tiene distribución de Haar sobre $\mathcal{U}(n)$.

Antes de proceder a probar la última afirmación, hay que hacer un par de observaciones.

Observaciones

- Para el caso simpléctico los elementos de \mathbb{H} son de la forma $h = \alpha_1 + i\alpha_2 + j\alpha_3 + k\alpha_4$ y se define h^* como

$$h^* = \alpha_1 - i\alpha_2 - j\alpha_3 - k\alpha_4 \in \mathbb{H},$$

del mismo modo para $\mathbf{a} = (a_1 \dots, a_n)^t \in \mathbb{H}^n$ y $\mathbf{b} = (b_1 \dots, b_n)^t \in \mathbb{H}^n$ se define

$$\mathbf{a}^* \mathbf{b} = \sum_{i=1}^n a_i^* b_i.$$

- Por la conmutatividad de los grupos ortogonal y unitario en tales casos el orden del producto no tiene relevancia, sin embargo en el caso simpléctico el producto no es conmutativo y habrá que tener especial cuidado.

Ahora bien para ver que la matriz \mathbf{U} definida en (2.2) tiene distribución de Haar sobre $\mathcal{U}(n)$ se probará que cumple la propiedad de invarianza bajo multiplicaciones por la izquierda de elementos de $\mathcal{U}(n)$.

Proposición 2.2. Sea $\mathbf{Z} = (z_{ij})_{1 \leq i, j \leq n}$ en $\text{GinUE}(n)$ cuya representación por columnas es

$$\mathbf{Z} = (\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n).$$

Si \mathbf{U} es la matriz definida en (2.2), entonces para toda $\mathbf{V} \in \mathcal{U}(n)$ las matrices \mathbf{U} y $\mathbf{V}\mathbf{U}$ tienen la misma distribución.

Demostración. Sea $\mathbf{V} \in \mathcal{U}(n)$, primero veamos que las matrices \mathbf{Z} y $\mathbf{V}\mathbf{Z}$ tienen la misma distribución. Para ello denotemos por ξ_{ij} a las entradas de $\mathbf{V}\mathbf{Z}$, es decir

$$\xi_{ij} = \sum_{k=1}^n v_{ik} z_{kj}.$$

Resulta claro que las variables ξ_{ij} siguen una distribución gaussiana y calculando directamente se obtiene

$$\mathbb{E}[\xi_{ij}] = \sum_{k=1}^n v_{ik} \mathbb{E}[z_{kj}] = 0$$

$$\mathbb{E}[\xi_{ij}\xi_{ij}^-] = \sum_{k=1}^n |v_{ik}|^2 \mathbb{E}[z_{kj}z_{kj}^-] = \sum_{k=1}^n |v_{ik}|^2 = 1,$$

y

$$\begin{aligned} \mathbb{E}[\xi_{ij}\xi_{sr}^-] &= \mathbb{E}\left[\left(\sum_{k=1}^n v_{ik}z_{kj}\right)\left(\sum_{m=1}^n v_{sm}^-z_{mr}^- \right)\right] \\ &= \sum_{k=1}^n \sum_{m=1}^n v_{ik}\bar{v}_{sm} \mathbb{E}[z_{kj}\bar{z}_{mr}] = \delta_{jr} \sum_{k=1}^n v_{ik}\bar{v}_{sk} = \delta_{jr}\delta_{is}. \end{aligned}$$

por lo que afirmamos que las variables ξ_{ij} son i.i.d. con distribución gaussiana $\mathbb{N}\mathbb{F}_2(0, 1)$ y por tanto

$$\mathbf{Z} \stackrel{\mathcal{L}}{=} \mathbf{V}\mathbf{Z}.$$

Ahora bien, la i -ésima columna de $\mathbf{V}\mathbf{U}$ es precisamente $\mathbf{V}\mathbf{u}_i$, que puede escribirse como

$$\begin{aligned} \mathbf{V}\mathbf{u}_i &= \frac{\mathbf{V}\mathbf{z}_i - \sum_{k=1}^{j-1} \mathbf{V}\mathbf{u}_k \cdot (\mathbf{u}_k^* \mathbf{z}_i)}{\|\mathbf{V}\mathbf{z}_i - \sum_{k=1}^{j-1} \mathbf{V}\mathbf{u}_k \cdot (\mathbf{u}_k^* \mathbf{z}_i)\|} \\ &= \frac{\mathbf{V}\mathbf{z}_i - \sum_{k=1}^{j-1} \mathbf{V}\mathbf{u}_k \cdot ((\mathbf{V}\mathbf{u}_k)^* \mathbf{V}\mathbf{z}_i)}{\|\mathbf{V}\mathbf{z}_i - \sum_{k=1}^{j-1} \mathbf{V}\mathbf{u}_k \cdot ((\mathbf{V}\mathbf{u}_k)^* \mathbf{V}\mathbf{z}_i)\|}. \end{aligned}$$

De la última expresión es claro que los vectores columna de $\mathbf{V}\mathbf{U}$ son la ortonormalización de Gram-Schmidt de los vectores columna de la matriz $\mathbf{V}\mathbf{Z}$. Como ya se ha probado que \mathbf{Z} y $\mathbf{V}\mathbf{Z}$ tienen la misma distribución se puede concluir que \mathbf{U} y $\mathbf{V}\mathbf{U}$ también. ■

De este resultado y el hecho de que la invarianza izquierda basta para caracterizar a la medida de Haar sobre un grupo compacto se concluye que la matriz \mathbf{U} , obtenida a través del proceso de ortonormalización de Gram-Schmidt, tiene distribución de Haar sobre $\mathcal{U}(n)$.

Una de las desventajas de utilizar el Proceso de Ortonormalización de Gram-Schmidt es su inestabilidad numérica debido a la acumulación de errores por redondeo, razón por la que pocas veces es implementado en la práctica.

Con el fin de solucionar esta deficiencia, se han propuesto varias alternativas al algoritmo clásico de Gram-Schmidt. Por ejemplo, en 1967 Rutishauser [59] propuso una modificación que consiste en una reorganización de los cálculos. La idea es tratar a los vectores de forma simultánea y no secuencial, con

lo que se evitan dificultades y que conduce a un algoritmo numérico robusto. Entre otras propuestas notorias pueden mencionarse el Proceso de Gram-Schmidt Modificado y el Proceso de Gram-Schmidt Iterativo, ambos numéricamente estables. Se han hecho diversos estudios y comparaciones, en términos de eficiencia y exactitud, entre estas y otros métodos alternativos con el algoritmo clásico.

A pesar del desarrollo de las modificaciones para salvar el problema de inestabilidad, el uso del Algoritmo de Ortonormalización de Gram-Schmidt para encontrar la descomposición QR no se hizo muy popular. Una de las razones es que por la misma época en que surgieron las modificaciones Businger y Golub [8] propusieron una nueva forma para calcular la descomposición. Su método consiste en el uso de reflexiones Householder y mostró ser estable numéricamente, así como eficiente en el sentido de número de operaciones que realiza. En la siguiente sección se presentara este método con detalle.

2.1.2. Reflexiones Householder

Debido a su eficiencia y estabilidad numérica el método de Reflexiones Householder es uno de los más usados en la práctica. Por simplicidad nos referiremos al caso ortogonal, los casos unitario y simpléctico tienen un tratamiento análogo.

Definición 10. Sea $\mathbf{v} \in \mathbb{R}^n - \{0\}$, la matriz definida de la siguiente forma

$$\mathbf{H} = \mathbf{I} - \frac{2}{\mathbf{v}^t \mathbf{v}} \mathbf{v} \mathbf{v}^t,$$

se denomina transformación, reflexión, o matrix Householder, correspondiente al vector Householder \mathbf{v} . Note que la acción de una reflexión Householder sobre un vector $\mathbf{x} \in \mathbb{R}^n$ consiste en reflejar \mathbf{x} en el hiperplano generado por $\{\mathbf{v}\}^\perp$.

De la definición es fácil ver que las matrices Householder son simétricas y ortogonales. La siguiente propiedad resultará de importancia para obtener la descomposición QR.

Proposición 2.3. Para todo vector $\mathbf{x} = (x_1, \dots, x_n)^t$ en \mathbb{R}^n existe una reflexión Householder $\mathbf{H}_\mathbf{x}$ que satisface

$$\mathbf{H}_\mathbf{x} \mathbf{x} = r \mathbf{e}_1,$$

donde $\mathbf{e}_1 = (1, 0, \dots, 0)^t$ y $r = \pm \|\mathbf{x}\|_2$.

Demostración. Se desea que $\mathbf{H}_\mathbf{x} \mathbf{x}$ pertenezca al subespacio generado por $\{\mathbf{e}_1\}$, de la definición de reflexión Householder se tiene

$$\mathbf{H}_\mathbf{x} \mathbf{x} = \mathbf{x} - \frac{2}{\mathbf{v}^t \mathbf{v}} \mathbf{v}^t \mathbf{x} \mathbf{v},$$

de acuerdo a esta última expresión, \mathbf{v} debe pertenecer al subespacio generado por $\{\mathbf{e}_1, \mathbf{x}\}$, esto es

$$\mathbf{v} = \mathbf{x} + \alpha \mathbf{e}_1,$$

entonces

$$\mathbf{v}^t \mathbf{x} = \mathbf{x}^t \mathbf{x} + \alpha \mathbf{x}_1$$

$$\mathbf{v}^t \mathbf{v} = \mathbf{x}^t \mathbf{x} + 2\alpha \mathbf{x}_1 + \alpha^2.$$

De lo anterior se sigue que

$$\mathbf{H}_x \mathbf{x} = \left(1 - 2 \frac{\mathbf{x}^t \mathbf{x} + \alpha \mathbf{x}_1}{\mathbf{x}^t \mathbf{x} + 2\alpha \mathbf{x}_1 + \alpha^2} \right) \mathbf{x} - 2\alpha \frac{\mathbf{v}^t \mathbf{x}}{\mathbf{v}^t \mathbf{v}} \mathbf{e}_1,$$

considerando $\alpha = \pm \|\mathbf{x}\|_2$ en esta última expresión, se tiene que

$$\mathbf{v} = \mathbf{x} \pm \|\mathbf{x}\|_2 \mathbf{e}_1,$$

y

$$\mathbf{H}_x \mathbf{x} = r \mathbf{e}_1,$$

tal como se quería demostrar. ■

Ahora veamos como obtener la descomposición QR utilizando transformaciones de tipo Householder.

Método de Reflecciones Householder

- Considere una matriz $\mathbf{Z} \in \text{GinOE}(n)$, cuyas columnas se denotarán por $\{\mathbf{z}_i\}_{i=1}^n$.
- Sea \mathbf{H}_1 la reflexión Householder que satisface que $\mathbf{H}_1 \mathbf{z}_1 = r_{11} \mathbf{e}_1$, donde $r_{11} = \pm \|\mathbf{z}_1\|^2$. Entonces

$$\mathbf{H}_1 \mathbf{Z} = \begin{pmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & & & \\ \vdots & & \mathbf{\Gamma}_1 & \\ 0 & & & \end{pmatrix}.$$

- Se continúa este proceso ahora con la matriz $\mathbf{\Gamma}_1$.
- De esta forma se obtiene una sucesión de transformaciones Householder que satisfacen

$$\mathbf{H}_{n-1} \cdots \mathbf{H}_2 \mathbf{H}_1 \mathbf{Z} = \mathbf{R},$$

donde \mathbf{R} es triangular superior por construcción. Así que basta tomar

$$\mathbf{Q} = \mathbf{H}_1 \mathbf{H}_2 \cdots \mathbf{H}_{n-1},$$

para obtener la descomposición $\mathbf{Z} = \mathbf{QR}$ con \mathbf{Q} una matriz ortogonal y \mathbf{R} triangular superior.

2.1.3. Observaciones sobre la Descomposición QR

De acuerdo al material presentado en la sección previa, después de elegir un método para encontrar la descomposición QR, generar una matriz aleatoria con distribución de Haar sería una tarea directa. Sin embargo, es necesario hacer ciertas observaciones al respecto.

En primer lugar, la descomposición $Z = QR$ no es única si no se le piden condiciones extras. Para verlo basta considerar una matriz diagonal D de la forma

$$D = \begin{pmatrix} e^{i\theta_1} & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 & \dots & e^{i\theta_n} \end{pmatrix} = \text{diag}(e^{i\theta_1}, \dots, e^{i\theta_n}),$$

entonces las matrices

$$Q' = QD \quad \text{y} \quad R' = D^{-1}R$$

son también unitaria y triangular superior, respectivamente; y se tiene que

$$Z = QR = Q'R',$$

con lo que tenemos una nueva descomposición de Z .

Para tener una descomposición única es suficiente pedir que se cumpla la condición de que los elementos de la diagonal de R sean positivos. Esto puede lograrse definiendo la matriz diagonal

$$\Lambda = \text{diag} \left(\frac{r_{11}}{\|r_{11}\|}, \frac{r_{22}}{\|r_{22}\|}, \dots, \frac{r_{nn}}{\|r_{nn}\|} \right),$$

y considerar las nuevas matrices

$$\tilde{Q} = Q\Lambda \quad \text{y} \quad \tilde{R} = \Lambda^{-1}R,$$

que conforman la descomposición única

$$Z = \tilde{Q}\tilde{R}$$

con R triangular superior y \tilde{Q} unitaria y con distribución de Haar sobre $\mathcal{U}(n)$.

La importancia de tener una descomposición QR única radica en que se ha observado que si no se pide tal condición al implementar los métodos para generar matrices aleatorias con distribución de Haar que involucran esta descomposición, las matrices que se obtienen como resultado no cumplen con las propiedades que se esperan de una matriz con distribución de Haar. Esto se debe a que existen problemas numéricos con las funciones de descomposición QR que introducen un sesgo provocando que

las matrices Q resultantes casi no tengan valores propios cercanos a uno.

Con el fin de ilustrar este punto, se implementó una rutina usando Mathematica 6.0, donde se reprodujeron los experimentos hechos por Edelman y Rao [23] en 2005, los cuales consisten en generar matrices aleatorias con entradas complejas independientes y con distribución gaussiana Z , obtener una descomposición QR de Z , encontrar los ángulos correspondientes a los eigenvalores de las matrices Q resultantes y graficar su histograma. Bajo el supuesto de que las matrices Q tienen distribución de Haar sobre $\mathcal{U}(n)$, la distribución de los eigenvalores debería ser uniforme, por lo que se espera que los histogramas de los ángulos correspondientes sea similar al de una densidad uniforme en -2π y 2π . Posteriormente se repite este proceso, ahora utilizando la descomposición QR única de Z . La rutina para el primer experimento es la siguiente:

```
Needs["Histograms`"]
RR := RandomReal[NormalDistribution[0, 1]]
RC := RR/Sqrt[2] + I*(RR/Sqrt[2])
RG[n_] := Table[RC, {n}, {n}];
n = 50;
m = 1000;
k = n*m;
anglesvector = Table[0, {k}];
For[i = 1, i < m + 1,
  Z = RG[n];
  U = QRDecomposition[Z][[1]];
  values = Eigenvalues[U];
  angles = Arg[values];
  For[j = 1, j < n + 1,
    anglesvector[[j + (i - 1)*n]] = angles[[j]];
    j++]
  i++]
Histogram[anglesvector];
```

Al ejecutar esta rutina se produjeron 10,000 matrices aleatorias con distribución de Haar sobre $\mathcal{U}(50)$ y se obtuvo el histograma de los ángulos correspondientes a sus eigenvalores, mismo que se presenta en la Figura 2.1.

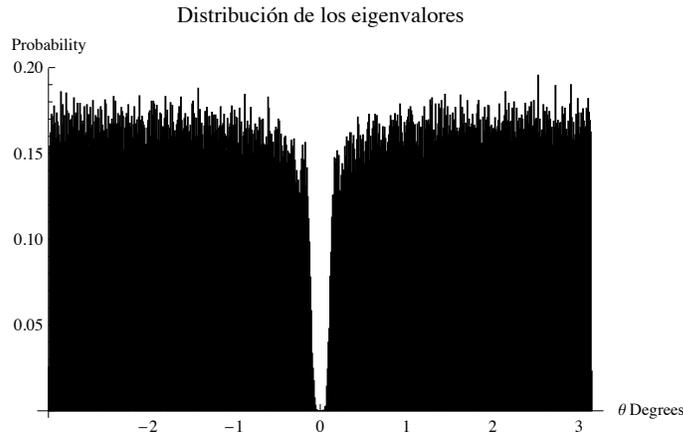


Figura 2.1: Histograma de los ángulos correspondientes a los eigenvalores de 10000 matrices de 50×50 generadas usando una descomposición QR.

Por otra parte, el algoritmo modificado para obtener la descomposición única se presenta a continuación.

```
Needs["Histograms`"]
RR := RandomReal[NormalDistribution[0, 1]]
RC := RR/Sqrt[2] + I*(RR/Sqrt[2])
RG[n_] := Table[RC, {n}, {n}];
n = 50;
m = 5000;
k = n*m;
anglesvector = Table[0, {k}];
For[i = 1, i < m + 1,
  Z = RG[n];
  U = QRDecomposition[Z][[1]];
  R = QRDecomposition[Z][[2]];
  U = U*DiagonalMatrix[Sign[Diagonal[R]]];
  values = Eigenvalues[U];
  angles = Arg[values];
  For[j = 1, j < n + 1,
    anglesvector[[j + (i - 1)*n]] = angles[[j]];
    j++]
  i++]
Histogram[anglesvector];
```

Con este algoritmo se obtiene el histograma de la Figura 2.2.

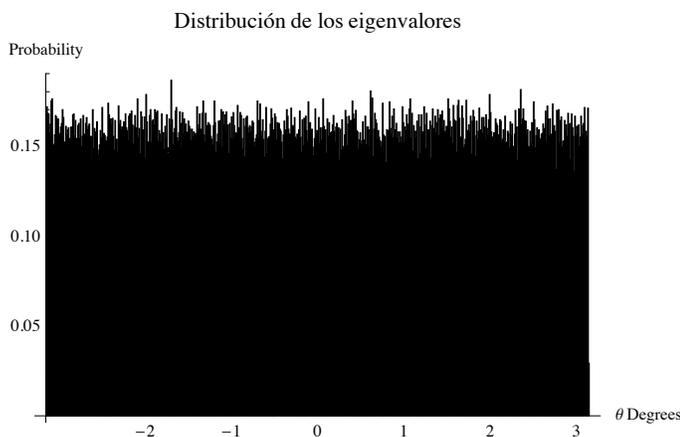


Figura 2.2: Histograma de los ángulos correspondientes a los eigenvalores de 10000 matrices de 50×50 generadas usando la descomposición QR única.

Recientemente Mezzadri [51] presentó una revisión de los métodos de generación de matrices aleatorias en los grupos compactos clásicos. En su artículo Mezzadri expone una discusión del problema que antes mencionamos y realiza experimentos similares a los de Edelman y Rao usando Phyton y otros paquetes de programación, notando que en todos los casos se observó el mismo comportamiento.

2.2. Una Interpretación Algebraica

De acuerdo a lo que se ha discutido antes es posible generar matrices aleatorias en los grupos compactos clásicos, $\mathcal{O}(n)$, $\mathcal{U}(n)$ y $\mathcal{S}p(n)$ usando poco más que métodos de álgebra lineal. A pesar de su simplicidad, estos algoritmos son casos particulares de un método general para generar elementos de grupos finitos. Tal método fue expuesto por primera vez en el trabajo de Diaconis y Shahshahani [27] y fue denominado Algoritmo del Subgrupo.

Algoritmo del Subgrupo

La idea general de este algoritmo es la siguiente, dado G un grupo finito del que se desea generar elementos $g \in G$ en forma aleatoria. Bajo el supuesto de que existe una cadena anidada de subgrupos de G , es decir subgrupos $G_i \subset G$ tales que

$$G_0 \subset G_1 \subset \cdots \subset G_{r-1} \subset G_r = G,$$

considere C_i las clases laterales de G_i en G_{i+1} respectivamente. Ésto induce una representación de G

como

$$G = C_{r-1} \times \cdots \times C_1 \times C_0,$$

de lo cual se sigue que cualquier elemento $g \in G$ tendrá una única representación de la forma

$$g = g_0 g_1 \cdots g_{r-1}, \quad g_i \in C_i.$$

Así que para generar un elemento $g \in G$ uniformemente, basta elegir uniformemente elementos g_i en C_i y tomar el producto.

Ejemplo 1: El grupo de permutaciones de n elementos.

Para ilustrar el algoritmo del subgrupo considere el problema de generar permutaciones aleatorias de n letras. Considere $G = S_n$ y la cadena de subgrupos

$$S_1 \subset S_2 \subset \cdots \subset S_{n-1} \subset S_n, \quad (2.3)$$

donde $S_{n-i}, i = 1, \dots, n$, representa el conjunto de todas las permutaciones de n elementos con los i primeros elementos fijos. En este ejemplo se observa claramente que es posible que existan diferentes cadenas de subgrupos, que inducirán diferentes representaciones del mismo grupo, por ejemplo considere ahora la cadena

$$S_1 \subset \cdots \subset S_{n-4} \subset S_{n-2} \subset S_n, \quad (2.4)$$

o incluso la cadena trivial

$$S_1 \subset S_n, \quad (2.5)$$

de hecho existen $n!$ posibles elecciones de cadenas de subgrupos para este grupo, sin embargo el método funciona independientemente de las posibles representaciones.

Ejemplo 2: El grupo de matrices Ortogonales.

Ahora tratemos el caso del grupo de matrices aleatorias ortogonales de tamaño $n \times n$, i.e. $G = \mathcal{O}(n)$; para utilizar el Algoritmo del Subgrupo puede considerarse la cadena de subgrupos

$$\mathcal{O}(1) \subset \mathcal{O}(2) \subset \cdots \subset \mathcal{O}(n-2) \subset \mathcal{O}(n-1) \subset \mathcal{O}(n) \quad (2.6)$$

donde $\mathcal{O}(n-1)$ es el subgrupo de elementos de $\mathcal{O}(n)$ fijando su primera columna como el vector canónico e_1 . Así, si deaseamos generar un elemento aleatorio con distribución uniforme sobre $\mathcal{O}(n)$ de acuerdo al Algoritmo del Subgrupo tomando la cadena (2.6), se tiene la representación

$$\mathcal{O}(n) = \mathcal{O}(1) \times \mathcal{O}(2)/\mathcal{O}(1) \times \cdots \times \mathcal{O}(n)/\mathcal{O}(n-1),$$

por lo que si sólo consideramos la parte superior de la cadena, es decir

$$\mathcal{O}(n-1) \subset \mathcal{O}(n)$$

bastará generar elementos aleatorios con distribución uniforme sobre el cociente $\mathcal{O}(n)/\mathcal{O}(n-1)$ y sobre el subgrupo $\mathcal{O}(n-1)$.

Aquí es importante hacer un par de observaciones: en primer lugar el grupo $\mathcal{O}(n-1)$ puede verse como subgrupo de $\mathcal{O}(n)$ al fijar el vector canónico e_1 . Además los elementos del cociente $\mathcal{O}(n)/\mathcal{O}(n-1)$ pueden definirse completamente diciendo a donde va e_1 , es decir podemos identificar este cociente con \mathbb{S}^{n-1} la esfera unitaria de dimensión $n-1$ y como dado cualquier vector $\mathbf{x} \in \mathbb{S}^{n-1}$ podemos considerar su reflexión Householder asociada, es decir,

$$\mathbf{I} - 2\mathbf{x}^t\mathbf{x},$$

entonces para todo $\mathbf{v} \in \mathbb{S}^n$ la reflexión anterior con \mathbf{x} definido como

$$\mathbf{x} = \frac{\mathbf{e}_1 - \mathbf{v}}{c} \quad \text{con} \quad c = \sqrt{(\mathbf{e}_1 - \mathbf{v})(\mathbf{e}_1 - \mathbf{v})'},$$

transformará al vector \mathbf{e}_1 en \mathbf{v} . Con ésto se tiene definido como generar elementos aleatorios uniformes en el cociente de $\mathcal{O}(n)$ en $\mathcal{O}(n-1)$, así que sólo resta generar un elemento aleatorio uniformemente sobre $\mathcal{O}(n-1)$, con lo que el problema se ha vuelto inductivo.

Capítulo 3

Aproximación de Matrices

Aleatorias de Haar

En este capítulo discutiremos varios resultados sobre dos enfoques de aproximación de matrices aleatorias con distribución de Haar sobre los grupos compactos clásicos matrices de entradas independiente y con distribución gaussiana estándar.

La primera consiste en estudiar la “cercanía” de estas matrices por medio de la distancia variacional entre sus distribuciones conjuntas. Recordemos que dadas μ y ν medidas de probabilidad sobre (Ω, \mathcal{F}) , la distancia variacional entre ellas, a la que denotaremos por $\rho(\mu, \nu)$, se define como

$$\rho(\mu, \nu) := 2 \sup_{A \in \mathcal{F}} |\mu(A) - \nu(A)|,$$

en particular cuando $(\Omega, \mathcal{F}) = (\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ y sabemos que μ y ν son absolutamente continuas con respecto a la medida de Lebesgue y tienen densidades $f(\mathbf{x})$ y $g(\mathbf{x})$ respectivamente, la distancia variacional entre ellas queda dada por

$$\rho(\mu, \nu) = \int_{\mathbb{R}^n} |f(\mathbf{x}) - g(\mathbf{x})| d\mathbf{x}. \quad (3.1)$$

Por otra parte, el segundo enfoque de aproximación se refiere en estudiar el comportamiento límite del máximo de la norma de las diferencias entre sus entradas. Es decir, si consideramos $\Gamma_n = (\gamma_{ij})_{1 \leq i, j \leq n}$ y $\mathbf{Z}_n = (z_{ij})_{1 \leq i, j \leq n}$ matrices aleatorias, entonces será de interés el comportamiento siguiente cantidad

$$\epsilon(m_n) := \max_{1 \leq i \leq n, 1 \leq j \leq m} |\gamma_{ij} - z_{ij}|,$$

cuando n tiende a infinito.

Convenciones y notación

En lo sucesivo seguiremos las siguientes convenciones y notación

- Sea \mathbf{X} una matriz aleatoria, denotaremos por $\mathcal{L}(\mathbf{X})$ a la distribución conjunta de todas sus entradas.
- Siempre que nos refiramos al bloque superior izquierdo de tamaño $p \times q$ de una matriz dada, estaremos considerando la matriz formada por sus primeras p filas y q columnas.
- En particular con \mathbf{U}_n denotaremos al bloque superior izquierdo de tamaño $p_n \times q_n$ de alguna matriz aleatoria $\Gamma_n = (\gamma_{ij})$ con distribución de Haar sobre $\mathcal{G}_\beta(n)$, esto es

$$\Gamma_n = \begin{pmatrix} \mathbf{U}_n & * \\ * & * \end{pmatrix}.$$

3.1. Sobre un Problema Planteado por Diaconis

El resultado que enunciamos a continuación fue probado en 1992 Diaconis, Eaton y Lauritzen [15] y constituye el punto inicial del problema de aproximación que nos concierne.

Teorema 3.1. *Sea $\Gamma_n = (\gamma_{ij})$ una matriz aleatoria con distribución de Haar sobre $\mathcal{O}(n)$ con \mathbf{U}_n su bloque superior izquierdo de tamaño $p_n \times q_n$, \mathbf{Z}_n una matriz aleatoria de tamaño $p_n \times q_n$, con entradas i.i.d. gaussianas estándar y sea*

$$\delta_n = \rho(\mathcal{L}(\sqrt{n}\mathbf{U}_n), \mathcal{L}(\mathbf{Z}_n)),$$

la distancia variacional entre las distribuciones de $\sqrt{n}\mathbf{U}_n$ y \mathbf{Z}_n . Si $p_n = o(n^\alpha)$ y $q_n = o(n^\alpha)$ con $\alpha = \frac{1}{3}$, entonces

$$\lim_{n \rightarrow \infty} \delta_n = 0.$$

El Teorema anterior tiene precedentes en las ideas propuestas en el trabajo de Borel [7] en 1906, de las que se sigue la siguiente convergencia

$$P(\sqrt{n}\gamma_{11} \leq x) \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left\{-\frac{t^2}{2}\right\} dt, \quad x \in \mathbb{R}$$

cuando $n \rightarrow \infty$. Esto es, para n suficientemente grande la distribución de cada entrada de la matriz Γ_n , escalada por \sqrt{n} , se asemeja a la de una variable aleatoria gaussiana estándar.

Posteriormente, Stam [63] probó que si consideramos el vector γ formado por las primeras m componentes de la primera columna de Γ_n escalada por \sqrt{n} , es decir

$$\gamma = \sqrt{n}(\gamma_{11}, \dots, \gamma_{m1})'$$

y el vector

$$\mathbf{z} = (z_{11}, \dots, z_{m1})'$$

de entradas i.i.d. gaussianas estándar, entonces la distancia variacional entre sus distribuciones tiende a cero cuando $n \rightarrow \infty$, siempre que $m = o(\sqrt{n})$. En 1987 Diaconis y Freedman [13] probaron que esta convergencia sigue cumpliéndose bajo la condición $m = o(n)$. Esta convergencia nos asegura que para n suficientemente grande las subcolumnas de la matriz Γ_n se comportan como vectores de entradas gaussianas e independientes.

Siguiendo el orden de estas ideas, es natural preguntarnos por el comportamiento de la distribución de un bloque de la matriz Γ_n . Consideremos U_n el bloque de tamaño $p_n \times q_n$ de Γ_n . De acuerdo a los resultados precedentes, es natural pensar que cuando n crece y se cumplen ciertas condiciones sobre p_n y q_n , la distribución de $\sqrt{n}U_n$ será cercana, en algún sentido, a la distribución de una matriz Z_n de tamaño $p_n \times q_n$ con entradas i.i.d. con distribución gaussiana estándar. El Teorema 3.1 dio condiciones suficientes sobre el orden de p_n y q_n para la convergencia de la distancia variacional de las distribuciones conjuntas de $\sqrt{n}U_n$ y Z_n tienda a cero cuando n crece a infinito.

Posteriormente, Diaconis [17] planteó formalmente el problema de encontrar el máximo orden de las dimensiones del bloque, de tal forma que la distribución conjunta de $\sqrt{n}U_n$ se aproxime, en algún sentido (e.g. distancia variacional) a la distribución conjunta de una matriz aleatoria Z_n de tamaño $p_n \times q_n$ con entradas i.i.d. con distribución gaussiana estándar, cuando n tiende a infinito.

Después de la publicación del Teorema 3.1 surgieron diversas conjeturas sobre el máximo orden de p_n y q_n para el que la convergencia $\delta_n \rightarrow 0$ siguiera cumpliéndose. En 2003 Collins [9] probó que $p_n = O(n^{\frac{1}{3}})$ y $q_n = O(n^{\frac{1}{3}})$ es suficiente para que la convergencia en distancia variacional sigue siendo válida.

Finalmente, en 2006 Jiang [37] demostró que el resultado de Diaconis (3.1) es válido para $\alpha = \frac{1}{2}$, es decir para cuando los bloques son del tamaño $p_n \times q_n$ con $p_n = o(n^{\frac{1}{2}})$ y $q_n = o(n^{\frac{1}{2}})$. Además probó que estos órdenes son los máximos tales que la distancia variacional δ_n , converge a cero cuando $n \rightarrow \infty$, dando con ello una primera respuesta al problema propuesto por Diaconis. En la sección siguiente se expone detalladamente esta respuesta.

3.2. Solución de Jiang

Los resultados que enunciamos a continuación muestran que el máximo α para el cual se cumple el Teorema 3.1 es $\frac{1}{2}$. Es importante mencionar que se cumplen para matrices aleatorias con distribución de Haar en $\mathcal{G}_\beta(n)$ para $\beta = 1, 2$, es decir en el grupo ortogonal y en el grupo unitario. También hay que decir que ambos casos fueron probados por Jiang [37, 39] el primero en 2006 y el segundo en 2009.

Las demostraciones siguen la misma idea y hacen uso de técnicas similares, por lo que sin perder generalidad aquí presentamos el caso unitario.

Teorema 3.2. *Sea Γ_n una matriz aleatoria con distribución de Haar en $\mathcal{U}(n)$ con \mathbf{U}_n su bloque superior izquierdo de tamaño $p_n \times q_n$, \mathbf{Z}_n una matriz aleatoria de tamaño $p_n \times q_n$, con entradas gaussianas independientes y sea δ_n la distancia variacional entre las distribuciones conjuntas de las matrices $\sqrt{n}\mathbf{U}_n$ y \mathbf{Z}_n , esto es*

$$\delta_n = \rho(\mathcal{L}(\sqrt{n}\mathbf{U}_n), \mathcal{L}(\mathbf{Z}_n)).$$

Si $p_n = o(\sqrt{n})$ y $q_n = o(\sqrt{n})$, entonces $\delta_n \rightarrow 0$, cuando $n \rightarrow \infty$.

Teorema 3.3. *Dados $x, y > 0$ para cada entero n considere*

$$p_n = \lfloor x\sqrt{n} \rfloor, \quad q_n = \lfloor y\sqrt{n} \rfloor,$$

\mathbf{U}_n el bloque superior izquierdo de tamaño $p_n \times q_n$, de Γ_n una matriz aleatoria con distribución de Haar en $\mathcal{U}(n)$ y \mathbf{Z}_n una matriz aleatoria de tamaño $p_n \times q_n$, con entradas gaussianas e independientes, entonces

$$\liminf_{n \rightarrow \infty} \delta_n \geq 2\Phi\left(\frac{xy}{2}\right) - 1, \quad (3.2)$$

donde $\delta_n = \rho(\mathcal{L}(\sqrt{n}\mathbf{U}), \mathcal{L}(\mathbf{Z}))$, y $\Phi(x)$ es la función de distribución de una variable aleatoria con distribución gaussiana estándar.

Antes de proseguir con las demostraciones de los teoremas anteriores, vale la pena hacer un análisis intuitivo de estos resultados. Como mencionamos antes, el problema planteado por Diaconis se refiere a encontrar el máximo orden de p_n y q_n para poder aproximar en algún sentido matrices aleatorias de Haar a través de matrices con entradas i.i.d. con distribución gaussiana. Los Teoremas 3.2 y 3.3 prueban para la aproximación en el sentido de distancia variacional el orden máximo para p_n y q_n es $o(\sqrt{n})$. Intuitivamente ¿porqué tiene sentido esta idea?

Diaconis y Freedman [13] demostraron que la distribución conjunta de las $m = o(n)$ primeras entradas de la primera columna de Γ_n podían aproximarse en el sentido de distancia variacional mediante variables aleatorias i.i.d. con distribución gaussiana estándar.

Si escribimos esta idea en términos del Teorema 3.2, tenemos que si \mathbf{U}_n es el bloque superior izquierdo, de tamaño $p_n \times 1$, de Γ_n y \mathbf{Z}_n es una matriz del mismo tamaño pero con entradas i.i.d. gaussianas, entonces la distancia variacional entre las distribuciones de $\sqrt{n}\mathbf{U}_n$ y \mathbf{Z}_n tiende a cero cuando $n \rightarrow \infty$, siempre que $p_n = o(n)$.

Hay que notar que para cada n fijo las matrices $\sqrt{n}\mathbf{U}_n$ y \mathbf{Z}_n tienen $p_n q_n = o(n)$ entradas, informalmente ésto nos dice que $o(n)$ entradas de Γ_n pueden ser aproximadas por variables aleatorias i.i.d. con ley gaussiana estándar. De modo que al considerar bloques de dimensión $p_n = o(n^\alpha)$ y $q_n = o(n^\alpha)$, se

está pensando en aproximar $p_n q_n = o(n^{2\alpha})$, por lo que no es extraño pensar que el máximo orden para α es $\frac{1}{2}$.

Más formalmente, la idea para probar estos resultados es la siguiente: sean $f_n(x)$ y $g_n(x)$ las funciones de densidad de $\sqrt{n}\mathbf{U}_n$ y \mathbf{Z}_n respectivamente, entonces

$$\begin{aligned}\delta_n &= \rho(\mathcal{L}(\sqrt{n}\mathbf{U}_n), \mathcal{L}(\mathbf{Z}_n)) \\ &= \int_{\mathbb{R}^{2p_n q_n}} |f_n(x) - g_n(x)| dx \\ &= \int_{\mathbb{R}^{2p_n q_n}} \left| \frac{f_n(x)}{g_n(x)} - 1 \right| g_n(x) dx\end{aligned}$$

de aquí se probará que si tanto p_n como q_n son del orden de \sqrt{n} entonces el cociente $\frac{f_n(x)}{g_n(x)}$ converge en distribución a una ley lognormal cuando $n \rightarrow \infty$ y converge a cero cuando p_n y q_n son del orden $o(\sqrt{n})$.

Para poder dar la demostración formal de los Teoremas 3.2 y 3.3 es necesario que tengamos algunos resultados previos, que expondremos a continuación.

3.2.1. Resultados Auxiliares

El Lema siguiente nos proporciona la expresión explícita de la densidad conjunta de un bloque de tamaño $p \times q$ tomado de una matriz aleatoria con distribución de Haar sobre $\mathcal{G}_2(n)$.

Lema 3.4. *Sea Γ_n una matriz con distribución de Haar en $\mathcal{U}(n)$, con \mathbf{U}_n su bloque superior izquierdo de tamaño $p \times q$. Si $p + q \leq n$ y $q \leq p$, entonces la función de densidad conjunta de las entradas de \mathbf{U}_n está dada por la siguiente expresión,*

$$f(U) = C_{p,q} \det(\mathbf{I} - U^*U)^{n-p-q} \mathbf{1}_{(0,1]}(\lambda_{\max}(U^*U)), \quad (3.3)$$

donde

$$\frac{1}{C_{p,q}} = \pi^{pq} \prod_{i=1}^q \frac{(n-i-p)!}{(n-i)!}. \quad (3.4)$$

Demostración El Teorema 5.1 en [9] muestra que la ecuación (3.3) es válida para alguna constante $C_{p,q} > 0$. Así que sólo resta calcular el valor de dicha constante, para ello consideremos la función

$$g(A) = C_{p,q} \det(\mathbf{I}_q - A)^{n-p-q} \mathbf{1}_{(0,1]}(\lambda_{\max}(A)), \quad (3.5)$$

donde $A = (a_{ij})_{1 \leq i, j \leq q}$ es una matriz positiva definida y $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_q)$, con λ_i los eigenvalores de A . Notemos que para toda matriz unitaria \mathbf{W} se cumple que

$$g(\mathbf{W}^* \Lambda \mathbf{W}) = g(\Lambda).$$

Por otra parte sabemos que la función de densidad conjunta de los eigenvalores de UU^* está dada por la expresión

$$C_{p,q} \frac{\pi^{q(p+q-1)}}{\Gamma_q(p)\Gamma_q(q)} \prod_{i=1}^q \lambda_i^{p-q} (1-\lambda_i)^{n-p-q} \prod_{1 \leq i < j \leq q} (\lambda_i - \lambda_j)^2 \mathbf{1}_{\{0 \leq \lambda_1 < \dots < \lambda_q \leq 1\}}, \quad (3.6)$$

donde $\Gamma_q(p) = \pi^{\frac{q(q-1)}{2}} \prod_{i=1}^q \Gamma(p-i+1)$. Así que para obtener el valor de $C_{p,q}$ integraremos (3.6) sobre \mathbb{R}^q . Para ello apelamos al hecho de que para cualquier función simétrica $h(\lambda_1, \dots, \lambda_q)$ se cumple que

$$\int_{0 \leq \lambda_1 \leq \dots \leq \lambda_q \leq 1} h(\lambda_1, \dots, \lambda_q) d\lambda_1 \dots d\lambda_q = \frac{1}{q!} \int_0^1 \dots \int_0^1 h(\lambda_1, \dots, \lambda_q) d\lambda_1 \dots d\lambda_q,$$

así como a la identidad conocida como Integral de Selberg con $\alpha = p - q + 1$, $\beta = n - p - q + 1$ y $\gamma = 1$, esto es

$$\begin{aligned} & \int_0^1 \int_0^1 \dots \int_0^1 \prod_{i=1}^n \lambda_i^{p-q} (1-\lambda_i)^{n-p-q} \prod_{1 \leq j < k \leq n} |\lambda_j - \lambda_k|^2 d\lambda_1 d\lambda_2 \dots d\lambda_n \\ &= \prod_{l=0}^{n-1} \frac{\Gamma(l+2)\Gamma(p-q+1+l)\Gamma(n-p-q+1+l)}{\Gamma(2)\Gamma(-q+1+n-q+n+l)}. \end{aligned}$$

De esta forma obtenemos

$$\begin{aligned} 1 &= \frac{1}{q!} C_{p,q} \frac{\pi^{q(p+q-1)}}{\Gamma_q(p)\Gamma_q(q)} \prod_{l=0}^{q-1} \frac{\Gamma(l+2)\Gamma(l+p-q+1)\Gamma(n+l-p-q+1)}{\Gamma(n-q+l+1)} \\ &= \frac{1}{q!} C_{p,q} \pi^{pq} \left(\frac{\prod_{i=1}^q \Gamma(i+1)}{\prod_{i=1}^q \Gamma(q-i+1)} \right) \left(\frac{\prod_{i=1}^q \Gamma(i+p-q)}{\prod_{i=1}^q \Gamma(p-i+1)} \right) \left(\frac{\prod_{i=1}^q \Gamma(n+i-p-q)}{\prod_{i=1}^q \Gamma(n-q+i)} \right) \\ &= \frac{1}{q!} C_{p,q} \pi^{pq} \Gamma(q+1) \prod_{i=1}^q \frac{\Gamma(n-p-i+1)}{\Gamma(n-q+i)}, \end{aligned}$$

por lo tanto

$$C_{p,q} = \frac{q!}{\pi^{pq} \Gamma(q+1)} \prod_{i=1}^q \frac{\Gamma(n-q+i)}{\Gamma(n-p-i+1)}$$

y con esto la prueba está completa. ■

Ahora que ya tenemos la densidad del bloque de la matriz aleatoria de Haar, podemos proceder a calcular la distancia variacional entre la distribución de este bloque escalado por \sqrt{n} y la de una matriz con entradas i.i.d. gaussianas. El siguiente resultado nos da una expresión para dicha distancia variacional.

Lema 3.5. Sean \mathbf{U}_n y \mathbf{Z}_n matrices aleatorias de tamaño $p_n \times q_n$ definidas como en el Teorema 3.2 y tales que $q_n \leq p_n$ y $p_n + q_n \leq n$. Si denotamos como δ_n a la distancia variacional entre las distribuciones de $\sqrt{n}\mathbf{U}_n$ y \mathbf{Z}_n , entonces

$$\delta_n = E[|K_n L_n - 1|], \quad (3.7)$$

con

$$K_n = \frac{1}{n^{pq}} \prod_{i=1}^q \frac{(n-i)!}{(n-i-p)!}, \quad (3.8)$$

$$L_n = \left(\exp \left\{ \sum_{i=1}^q \lambda_i \right\} \right) \left(\prod_{i=1}^q \left(1 - \frac{\lambda_i}{n} \right) \right)^{n-p-q} \mathbf{1}_{(0,n]}(\max_{1 \leq i \leq q} \lambda_i)$$

y donde $\lambda_1, \dots, \lambda_q$ son los eigenvalores de $\mathbf{Z}_n^* \mathbf{Z}_n$.

Demostración: Sean $f_n(x)$ y $g_n(x)$ las densidades de $\sqrt{n}\mathbf{U}_n$ y \mathbf{Z}_n , respectivamente. Note que podemos utilizar el Lema anterior para obtener una expresión explícita de $f_n(x)$ como

$$\begin{aligned} f_n(\mathbf{X}) &= \frac{1}{n^{pq}} f\left(\frac{\mathbf{X}}{\sqrt{n}}\right) \\ &= \frac{C_{p,q}}{n^{pq}} \det\left(I - \frac{\mathbf{X}^* \mathbf{X}}{n}\right)^{n-p-q} \mathbf{1}_{(0,1]}(\lambda_{\max}\left(\frac{\mathbf{X}^* \mathbf{X}}{n}\right)) \\ &= \frac{C_{p,q}}{n^{pq}} \det\left(I - \frac{\mathbf{X}^* \mathbf{X}}{n}\right)^{n-p-q} \mathbf{1}_{(0,1]}(\lambda_{\max}(\mathbf{X}^* \mathbf{X})) \end{aligned}$$

donde f es la densidad de \mathbf{U}_n tal como en (3.4). Por otra parte, es claro que

$$g_n(\mathbf{X}) = \frac{1}{\sqrt{\pi}^{2pq}} \exp\{-tr(\mathbf{X}^* \mathbf{X})\}.$$

Por lo tanto

$$\begin{aligned} \delta_n &= \rho(\mathcal{L}(\sqrt{n}\mathbf{U}_n), \mathcal{L}(\mathbf{Z}_n)) = \int_{\mathbb{R}^{2p_n q_n}} |f_n(x) - g_n(x)| dx \\ &= \int_{\mathbb{R}^{2p_n q_n}} \left| \frac{f_n(x)}{g_n(x)} - 1 \right| g_n(x) dx = E_{\mathbf{Z}} \left[\left| \frac{f_n(\mathbf{X}_n)}{g_n(\mathbf{X}_n)} - 1 \right| \right] \\ &= E_{\mathbf{Z}} \left[\left| \frac{\frac{C_{p,q}}{n^{pq}} \det\left(I - \frac{\mathbf{X}^* \mathbf{X}}{n}\right)^{n-p-q} \mathbf{1}_{(0,1]}(\lambda_{\max}(\mathbf{X}^* \mathbf{X}))}{\frac{1}{\sqrt{\pi}^{2pq}} \exp\{-tr(\mathbf{X}^* \mathbf{X})\}} - 1 \right| \right], \end{aligned}$$

con \mathbf{X} una matriz de tamaño $p_n \times q_n$ de entradas i.i.d. con distribución gaussiana estándar.

Para terminar la demostración basta notar que

$$\begin{aligned} \det\left(I - \frac{\mathbf{X}^* \mathbf{X}}{n}\right)^{n-p-q} &= \left[\prod_{i=1}^q \left(1 - \frac{\lambda_i}{n}\right) \right]^{n-p-q} \\ tr(\mathbf{X}^* \mathbf{X}) &= \sum_{i=1}^q \lambda_i, \end{aligned}$$

con λ_i son los eigenvalores de $\mathbf{X}^*\mathbf{X}$. Utilizando estas dos igualdades obtenemos la igualdad (3.7). ■

Ahora que tenemos una expresión para la distancia variacional procederemos con la aproximación que buscamos. Primero observemos que el primer término de la expresión que se encuentra dentro de la esperanza en (3.7) tiene una parte determinística K_n y otra aleatoria L_n . Los Lemas siguientes nos proporcionan aproximaciones para K_n y L_n cuando n suficientemente grande.

Lema 3.6. (Aproximación de K_n) Sean $x > 0$, $y > 0$, considere $p = \lfloor xn^{\frac{1}{2}} \rfloor$, $q = \lfloor yn^{\frac{1}{2}} \rfloor$ y K_n como en (3.8). Para n suficientemente grande se cumple que

$$K_n = \exp \left\{ -\frac{p^2q + pq^2}{2n} - \frac{2x^3y + 2xy^3 + 3x^2y^2}{12} + O\left(\frac{1}{\sqrt{n}}\right) \right\}. \quad (3.9)$$

Demostración: Partimos directamente de la definición de K_n

$$K_n = \frac{1}{n^{pq}} \prod_{j=1}^q \frac{(n-j)!}{(n-j-p)!} = \frac{1}{n^{pq}} \prod_{j=1}^q \prod_{i=-1}^{p-2} (n-i-j) = \prod_{j=1}^q \prod_{i=0}^{p-1} \left(1 - \frac{i+j}{n}\right),$$

tomamos el logaritmo y obtenemos

$$\log(K_n) = \sum_{j=1}^q \sum_{i=0}^{p-1} \log\left(1 - \frac{i+j}{n}\right). \quad (3.10)$$

Por otra parte, de calcular el desarrollo en serie de Maclaurin de la función $\log(1-x)$ sabemos que

$$\left| \log(1-x) + x + \frac{x^2}{2} \right| \leq x^3, \quad (3.11)$$

para x suficientemente pequeño. Notemos que $p \sim x\sqrt{n}$ y $q_n \sim y\sqrt{n}$ por hipótesis, de esto podemos concluir que

$$\lim_{n \rightarrow \infty} \frac{p+q}{n} = 0,$$

por lo tanto podemos utilizar la desigualdad (3.11) para cada uno de los términos de la suma en (3.10).

De esta observación y la desigualdad del triángulo concluimos que

$$\left| \log(K_n) + \frac{1}{n} \sum_{j=1}^q \sum_{i=0}^{p-1} (i+j) + \frac{1}{2n^2} \sum_{j=1}^q \sum_{i=0}^{p-1} (i+j)^2 \right| \leq \frac{1}{n^3} \sum_{j=1}^q \sum_{i=0}^{p-1} (i+j)^3, \quad (3.12)$$

para n suficientemente grande.

Con el fin de reescribir la última desigualdad, consideremos las cantidades

$$H_n = \frac{1}{n} \sum_{j=1}^q \sum_{i=0}^{p-1} (i+j) \quad \text{y} \quad I_n = \frac{1}{2n^2} \sum_{j=1}^q \sum_{i=0}^{p-1} (i+j)^2.$$

Siguiendo esta notación y utilizando el hecho de que $(i+j)^3 \leq (p+q)(i+j)^2$ para $0 \leq i \leq p$ y $1 \leq j \leq q$, la desigualdad (3.12) se escribe como

$$|\log(K_n) + H_n + I_n| \leq \frac{C}{\sqrt{n}} I_n, \quad (3.13)$$

para n suficientemente grande, donde $C = C(x, y)$ es una constante positiva.

Ahora bien, calculando directamente tenemos que

$$H_n = \frac{1}{n} \sum_{j=1}^q \sum_{i=0}^{p-1} (i+j) = \frac{q}{n} \sum_{i=0}^{p-1} i + \frac{p}{n} \sum_{j=1}^q j = \frac{pq(p-1)}{2n} + \frac{pq(q+1)}{2n} = \frac{pq(p+q)}{2n}. \quad (3.14)$$

Del mismo modo

$$\begin{aligned} I_n &= \frac{1}{2n^2} \sum_{j=1}^q \sum_{i=0}^{p-1} (i^2 + 2ij + j^2) \\ &= \frac{q}{2n^2} \frac{p(p-1)(2p-1)}{6} + \frac{p}{2n^2} \frac{q(q+1)(2q+1)}{6} + \frac{2}{2n^2} \frac{p(p-1)}{2} \frac{q(q+1)}{2}, \end{aligned}$$

pero apelando al hecho de que $p = x\sqrt{n} + O(1)$ y $q = y\sqrt{n} + O(1)$, concluimos que

$$I_n = \frac{2x^3y + 2xy^3 + 3x^2y^2}{12} + O\left(\frac{1}{\sqrt{n}}\right), \quad (3.15)$$

para n suficientemente grande.

Sustituimos (3.14) y (3.15) en la expresión (3.13), entonces

$$\left| \log(K_n) + \frac{pq(p+q)}{2n} + \frac{2x^3y + 2xy^3 + 3x^2y^2}{12} + O\left(\frac{1}{\sqrt{n}}\right) \right| \leq \frac{C}{\sqrt{n}} I_n, \quad (3.16)$$

para n suficientemente grande. De esta última desigualdad podemos concluir el resultado deseado. ■

Lema 3.7. (Aproximación de L_n) Sean $x > 0$, $y > 0$, $p = [xn^{\frac{1}{2}}]$ y $q = [yn^{\frac{1}{2}}]$. Si L_n es como en el Lema 3.5, entonces $e^{-a_n} L_n$ converge débilmente a la ley de $e^{\sigma\xi}$, donde

$$a_n = \frac{p^2q + pq^2}{2n} + \frac{x^3y + xy^3}{6}, \quad \sigma = \frac{xy}{2},$$

y ξ es una variable aleatoria gaussiana estándar.

Demostración: Nuestro objetivo es aproximar $e^{-a_n} L_n$, pero primero notemos que

$$\begin{aligned} L_n &= \exp \left\{ \sum_{i=1}^q \lambda_i \right\} \left(\prod_{i=1}^q \left(1 - \frac{\lambda_i}{n} \right) \right)^{n-p-q} \mathbf{1}_{(0,n]}(\max_{1 \leq i \leq q} \lambda_i) \\ &= \exp \left\{ \sum_{i=1}^q \lambda_i + \sum_{i=1}^q (n-p-q) \log \left(1 - \frac{\lambda_i}{n} \right) \right\} \mathbf{1}_{(0,n]}(\max_{1 \leq i \leq q} \lambda_i) \\ &= \exp \left\{ \sum_{i=1}^q \lambda_i + (n-p-q) \log \left(1 - \frac{\lambda_i}{n} \right) \right\} \mathbf{1}_{(0,n]}(\max_{1 \leq i \leq q} \lambda_i). \end{aligned}$$

De modo que si consideramos la función

$$f(x) = \begin{cases} x + (n-p-q) \log(1 - \frac{x}{n}) & \text{si } 0 \leq x < n \\ -\infty & \text{si } x \geq n, \end{cases}$$

tenemos que

$$\ln(L_n) = \sum_{i=1}^q f(\lambda_i). \quad (3.17)$$

Utilizando la expansión en serie de Maclaurin de $\log(1-x)$, concluimos que $x \in (0, n)$ existe $\xi = \xi_x \in (0, x)$ tal que

$$\log\left(1 - \frac{x}{n}\right) = -\frac{x}{n} - \frac{x^2}{2n^2} - \frac{x^3}{3n^3} - \frac{x^4}{4} \frac{1}{(\xi - n)^4}.$$

Luego,

$$f(x) = \frac{(p+q)x}{n} - \frac{(n-p-q)x^2}{2n^2} - \frac{(n-p-q)x^3}{3n^3} + g_n(\xi) \frac{x^4}{n^3}, \quad (3.18)$$

para todo $x \in (0, n)$ y donde $g_n(x) = -\frac{(n-p-q)n^3}{4(x-n)^4}$. Además, es fácil ver que

$$\sup_{0 \leq x \leq \alpha n} |g_n(x)| \leq \frac{1}{(1-\alpha)^4},$$

para $\alpha \in (0, 1)$.

Por otra parte, $\mathbf{Z}_n^* \mathbf{Z}_n$ satisface las hipótesis del Teorema 1.6, entonces existe una constante $c(x, y) > 0$ tal que

$$\frac{\lambda_{\max}(\mathbf{Z}_n^* \mathbf{Z}_n)}{\sqrt{n}} \xrightarrow{P} c(x, y) \quad (3.19)$$

cuando $n \rightarrow \infty$. Definimos los conjuntos

$$\Omega_n := \{ \text{máx}_{1 \leq i \leq q} \lambda_i \leq \sqrt{n}(c(x+y) + 1) \},$$

y por (3.19) se tiene que $\lim_{n \rightarrow \infty} P(\Omega_n) = 1$. Más aún, observemos que en cada conjunto Ω_n se satisface (3.18), por construcción. Como consecuencia podemos sustituir (3.18) en (3.17) sobre los conjuntos Ω_n ,

$$\begin{aligned} \ln(L_n) &= \sum_{i=1}^q \left[\frac{(p+q)\lambda_i}{n} - \frac{(n-p-q)\lambda_i^2}{2n^2} - \frac{(n-p-q)\lambda_i^3}{3n^3} + g_n(\xi) \frac{\lambda_i^4}{n^3} \right] \\ &= \frac{(p+q)}{n} \sum_{i=1}^q \lambda_i - \frac{(n-p-q)}{2n^2} \sum_{i=1}^q \lambda_i^2 - \frac{(n-p-q)}{3n^3} \sum_{i=1}^q \lambda_i^3 + \frac{\tilde{g}_n(\xi)}{n^3} \sum_{i=1}^q \lambda_i^4 \\ &= \frac{(p+q)}{n} \text{tr}(\mathbf{Z}_n^* \mathbf{Z}_n) - \frac{(n-p-q)}{2n^2} \text{tr}((\mathbf{Z}_n^* \mathbf{Z}_n)^2) \\ &\quad - \frac{(n-p-q)}{3n^3} \text{tr}((\mathbf{Z}_n^* \mathbf{Z}_n)^3) + \frac{\tilde{g}_n(x)}{n^3} \text{tr}((\mathbf{Z}_n^* \mathbf{Z}_n)^4) \end{aligned} \quad (3.20)$$

para $|\tilde{g}_n| \in [0, 2)$ y n suficientemente grande. Notemos que $\text{tr}((\mathbf{Z}_n^* \mathbf{Z}_n)^k)$ es una variable aleatoria bien definida y no depende del conjunto Ω_n . Por el Lema A.2

$$\frac{\text{tr}((\mathbf{Z}_n^* \mathbf{Z}_n)^4)}{n^3} = \frac{\text{tr}((\mathbf{Z}_n^* \mathbf{Z}_n)^4) - E[\text{tr}((\mathbf{Z}_n^* \mathbf{Z}_n)^4)]}{n^3} + O\left(\frac{q^5}{n^3}\right) \rightarrow 0$$

en probabilidad cuando $n \rightarrow \infty$. Del mismo Lema, considerando $\eta = \frac{x}{y}$ se siguen las siguientes convergencias en probabilidad

$$\begin{aligned} \frac{(p+q)}{2n^2} \text{tr}((\mathbf{Z}_n^* \mathbf{Z}_n)^2) &= \frac{(p+q)q^3}{2n^2} \frac{\text{tr}((\mathbf{Z}_n^* \mathbf{Z}_n)^2)}{q^3} \rightarrow \frac{xy(x+y)^2}{2}, \\ \frac{(n-p-q)}{3n^3} \text{tr}((\mathbf{Z}_n^* \mathbf{Z}_n)^3) &\sim \frac{q^4}{3n^2} \frac{\text{tr}((\mathbf{Z}_n^* \mathbf{Z}_n)^3)}{q^4} \rightarrow \frac{xy(x^2+y^2+3xy)^2}{3}, \end{aligned}$$

y finalmente

$$\frac{\text{tr}((\mathbf{Z}_n^* \mathbf{Z}_n)^4)}{n^3} = \frac{q^5}{n^3} \frac{\text{tr}((\mathbf{Z}_n^* \mathbf{Z}_n)^4)}{q^5} \rightarrow 0,$$

cuando $n \rightarrow \infty$. Por lo anterior se sigue que

$$\frac{(n-p-q)}{3n^3} \text{tr}((\mathbf{Z}_n^* \mathbf{Z}_n)^3) - \frac{\text{tr}((\mathbf{Z}_n^* \mathbf{Z}_n)^4)}{n^3} \rightarrow \frac{x^3y + xy^3}{6}$$

en probabilidad, cuando $n \rightarrow \infty$. Considerando estas convergencias junto con la ecuación (3.20) vemos que si

$$R_n = \frac{p+q}{n} \text{tr}(\mathbf{Z}_n^* \mathbf{Z}_n) - \frac{1}{2n} \text{tr}((\mathbf{Z}_n^* \mathbf{Z}_n)^2), \quad (3.21)$$

entonces para probar el resultado que queremos basta probar que $R_n - \frac{p^2q + pq^2}{2n}$ converge débilmente a una ley gaussiana con media cero y varianza $\frac{x^2y^2}{4}$. Para verificarlo, primero notemos que el Lema C.3 implica que

$$\begin{aligned} \text{tr}((\mathbf{Z}_n^* \mathbf{Z}_n)^2) &= \sum_{j=1}^q \sum_{i=1}^p |z_{ij}|^4 + \sum_{j=1}^q \sum_{i \leq l=1}^p |z_{ij}|^2 |z_{lj}|^2 \\ &\quad + \sum_{i=1}^p \sum_{j \leq k=1}^q |z_{ij}|^2 |z_{ik}|^2 + \sum_{i \neq k, j \neq l} z_{ij} z_{kj} z_{kl} z_{il}, \end{aligned}$$

luego $E[\text{tr}((\mathbf{Z}_n^* \mathbf{Z}_n)^2)] = pq(E[|z_{11}|^4]) + qp(p-1) + pq(q-1) = pq(p+q)$. También sabemos que $E[\text{tr}(\mathbf{Z}_n^* \mathbf{Z}_n)] = pq$, por lo tanto

$$E[R_n] = \frac{pq(p+q)}{n} - \frac{pq(p+q)}{2n} = \frac{pq(p+q)}{2n}.$$

Sea $h_i = \text{tr}(\mathbf{Z}_n^* \mathbf{Z}_n)^i - E[\text{tr}(\mathbf{Z}_n^* \mathbf{Z}_n)^i]$ para $i = 1, 2$, entonces es suficiente verificar que

$$W_n := \frac{p+q}{n} h_1 - \frac{1}{2n} h_2$$

converge en distribución a una ley gaussiana $N(0, \sigma^2)$. Note que

$$\text{var}[h_1] = \text{var}[\text{tr}(\mathbf{Z}_n^* \mathbf{Z}_n)] = pq.$$

Por otro lado, del Lema A.3, $\frac{\text{var}[h_2]}{n^2}$ converge débilmente a una constante positiva. Del Teorema 1.1 de [4] se sigue que $\left(\frac{h_1}{\sqrt{\text{var}(h_1)}}, \frac{h_2}{\sqrt{\text{var}(h_2)}} \right)$ converge en distribución a la ley gaussiana de media cero. Así,

W_n converge en distribución a una ley gaussiana $N(0, \sigma^2)$ y al ser uniformemente integrable se tiene que $\lim_{n \rightarrow \infty} E[W_n^2] = \sigma^2$. Por otra parte

$$\text{var}[W_n] = \frac{(p+q)^2}{n^2} \text{var}[\text{tr}(\mathbf{Z}_n^* \mathbf{Z}_n)] - \frac{p+q}{n^2} \text{cov}[\text{tr}(\mathbf{Z}_n^* \mathbf{Z}_n), \text{tr}(\mathbf{Z}_n^* \mathbf{Z}_n)^2].$$

De lo anterior se sigue que

$$\sigma^2 = \lim_{n \rightarrow \infty} \text{var}[W_n] = \frac{x^2 y^2}{4}$$

y la prueba está completa. ■

3.2.2. Pruebas de los Resultados Principales

Prueba del Teorema 3.3 *Demostración:* Del Lema 3.5 tenemos que

$$\delta_n = E[|K_n L_n - 1|]$$

y de los Lemas 3.6 y 3.7 se sigue que $K_n L_n$ converge débilmente a $\exp\left(-\frac{x^2 y^2}{4} + \frac{xy}{2} \xi\right)$ cuando $n \rightarrow \infty$, el Lema de Fatou nos permite concluir el resultado. ■

Prueba del Teorema 3.2 *Demostración:* Sin pérdida de generalidad supongamos que $q_n = p_n$ para todo $n \geq 1$. Sabemos que

$$\delta_n = E|K_n L_n - 1|,$$

con K_n y L_n dados como en (3.8). Más aún, del Lema 3.6 tenemos que

$$K_n \sim \exp\left\{-\frac{p^2 q + p q^2}{2n}\right\}, \quad (3.22)$$

para n suficientemente grande. Por lo tanto para probar el resultado basta con demostrar la siguiente convergencia en probabilidad

$$e^{-\frac{(p^2 q + p q^2)}{2n}} L_n \rightarrow 1, \quad (3.23)$$

cuando $n \rightarrow \infty$.

Seguiremos la notación y el esquema de la prueba del Lema 3.7 con $x = y = 0$. Para verificar (3.23) hay que demostrar que $R_n - \frac{p^2 q + p q^2}{2n}$ converge en probabilidad a cero, cuando $n \rightarrow \infty$. Directamente de la definición tenemos que

$$R_n - \frac{p^2 q + p q^2}{2n} = \frac{p+q}{n} \text{tr}(\mathbf{X}_n^* \mathbf{X}_n) - \frac{1}{2n} \text{tr}((\mathbf{Z}_n^* \mathbf{Z}_n)^2) \frac{p^2 q + p q^2}{2n} - \frac{p^2 q + p q^2}{2n},$$

y ya se ha mostrado que $E[R_n] = \frac{(p^2q+pq^2)}{2n}$, por lo tanto es suficiente probar que $\frac{p+q}{n}h_1$ y $\frac{1}{n}h_2$ convergen a cero en probabilidad, cuando $n \rightarrow \infty$. Pero ambos hechos se obtienen de la desigualdad de Chebyshev y las siguientes convergencias, que ya se han probado

$$\begin{aligned} \text{var} \left[\frac{p+q}{n} h_1 \right] &= \lim_{n \rightarrow \infty} \frac{(p+q)^2 pq}{n^2} = 0, \\ \text{var} \left[\frac{h_2}{n} \right] &\sim \frac{17p^4}{n^2} \rightarrow 0. \end{aligned}$$

De lo anterior se sigue (3.23). Utilizando ese hecho y (3.22) se tiene que $K_n L_n$ converge a 1 en probabilidad, cuando $n \rightarrow \infty$ y que $E[K_n L_n] = 1$. De estas dos afirmaciones se sigue el resultado que deseabamos probar. ■

3.3. Convergencia del Máximo de Diferencias de Entradas

El propósito de esta Sección es exponer los resultados principales del trabajo de Jiang [38] en el que expone una segunda respuesta al problema planteado por Diaconis, detallado anteriormente, acerca de la aproximación de matrices con distribución de Haar por medio de matrices con entradas i.i.d. con distribución gaussiana. Este segundo enfoque consiste en considerar matrices $\Gamma_n = (\gamma_{ij})_{1 \leq i, j \leq n}$ con distribución de Haar sobre $\mathcal{O}(n)$, $\mathcal{U}(n)$ y $\mathcal{S}p(n)$ generadas utilizando el proceso de ortonormalización de Gram-Schmidt sobre matrices $\mathbf{Z}_n = (z_{ij})_{1 \leq i, j \leq n}$ del correspondiente ensamble de Ginibre (GinOE, GinUE y GinSE respectivamente) y estudiar el comportamiento del máximo de las diferencias entre sus correspondientes entradas, es decir

$$\epsilon_n(m) := \max_{1 \leq i \leq n, 1 \leq j \leq m} |\sqrt{n}\gamma_{ij} - z_{ij}|,$$

donde m depende de n , cuando n crece a infinito.

Debemos mencionar que tanto los resultados principales de esta sección, como sus demostraciones provienen del trabajo de Jiang, pero las pruebas se presentan con mayor detalle.

3.3.1. Resultado Principal

Teorema 3.8. *Para cada $n \geq 2$ existen matrices $\Gamma_n = (\gamma_{ij})_{1 \leq i, j \leq n}$ y $\mathbf{Z}_n = (z_{ij})_{1 \leq i, j \leq n}$, cuyas entradas son variables aleatorias definidas sobre el mismo espacio de probabilidad y tales que*

- i) *Las entradas de \mathbf{Z}_n son variables aleatorias i.i.d. con distribución*
- ii) *La matriz Γ_n tienen distribución de Haar sobre $\mathcal{G}_\beta(n), \mathbb{NF}_\beta(0, 1)$.*

Para $1 \leq m \leq n$ sea

$$\epsilon_n(m) = \max_{1 \leq i \leq n, 1 \leq j \leq m} |\sqrt{n}\gamma_{ij} - z_{ij}|,$$

entonces

$$\begin{aligned} P(\epsilon_n(m) \geq rs + t) &\leq 4me^{-\frac{\beta nr^2}{16}} + mn\chi_\beta(n, s)e^{-\frac{\beta s^2}{2}} \\ &\quad + 3mn\chi_\beta(n, t) \left(1 + \frac{t^2}{3(m + t\sqrt{n})}\right)^{-\frac{\beta n}{2}}, \end{aligned}$$

para todo $r \in (0, \frac{1}{4})$, $s > 0$, $t > 0$, y $1 \leq m \leq \frac{nr}{2}$, donde

$$\chi_\beta(n, s) = \begin{cases} \frac{1}{s}, & \beta = 1 \\ 1, & \beta = 2 \\ 8ns^2 + 1, & \beta = 4. \end{cases}$$

y β toma los valores 1,2,4 de acuerdo al grupo $\mathcal{G}_\beta(n)$ al que pertenece la matriz Γ_n .

Corolario 1. Considere las hipótesis del teorema anterior. Si $m_n = o(\frac{n}{\log n})$, entonces $\epsilon_n(m_n) \rightarrow 0$ en probabilidad cuando n tiende a infinito.

3.3.2. Resultados Auxiliares

Antes de proceder a la demostración del Teorema 3.8 es necesario dar algunos Lemas previos, así como introducir la siguiente notación: Dado $\mathbf{x} = (x_1, \dots, x_n)^t$ un vector con entradas en \mathbb{F}_β , de aquí en adelante usaremos $\|\mathbf{x}\|$ para denotar al máximo de los valores absolutos de las entradas de \mathbf{x} , es decir

$$\|\mathbf{x}\| := \max_{1 \leq i \leq n} |x_i|.$$

Los siguientes lemas establecen resultados sobre la distribución de columnas y entradas de matrices aleatorias con distribución de Haar.

Lema 3.9. Sea $\Gamma_n = (\gamma_1, \dots, \gamma_n)$ una matriz aleatoria con distribución de Haar sobre $\mathcal{O}(n)$ ($\beta = 1$), $\mathcal{U}(n)$ ($\beta = 2$) o $\mathcal{Sp}(n)$ ($\beta = 4$), cualquier fila o columna de Γ_n tiene la misma distribución que

- $\frac{1}{\|X_n^t\|}(\xi_{11}, \dots, \xi_{n1})^t$ para $\beta = 1$,
- $\frac{1}{\|X_n^2\|}(\xi_{11} + i\xi_{12}, \dots, \xi_{n1} + i\xi_{n2})^t$ para $\beta = 2$,
- $\frac{1}{\|X_n^4\|}(\xi_{11} + i\xi_{12} + j\xi_{13} + k\xi_{14}, \dots, \xi_{n1} + i\xi_{n2} + j\xi_{n3} + k\xi_{n4})^t$ para $\beta = 4$,

donde $\{\xi_{pq}\}_{1 \leq p \leq n, 1 \leq q \leq 4}$ son variables aleatorias independientes y con distribución gaussiana $\mathbb{N}\mathbb{F}_\beta(0, 1)$

y $\|X_n^\beta\|^2 = \sum_{1 \leq p \leq n, 1 \leq q \leq \beta} \xi_{pq}^2$, para $\beta = 1, 2, 4$.

Demostración Sin pérdida de generalidad se probará el resultado para las columnas γ_i , las filas requiere un tratamiento análogo.

Los casos ortogonal y unitario ($\beta = 1, 2$ respectivamente) se siguen inmediatamente de la Proposición 1.8, notando que $\gamma_i = \Gamma_n e_i$.

Ahora probemos el caso simpléctico, para ello considere $\mathbf{Y} = (y_1, \dots, y_n)$ una matriz de $n \times n$, cuyas entradas $(y_{ij})_{1 \leq i, j \leq n}$ son i.i.d. y tienen la misma distribución que $\xi_{11} + i\xi_{12} + j\xi_{13} + k\xi_{14}$. Se aplica proceso de ortonormalización de Gram-Schmidt, por lo que se define

$$\mathbf{w}_1 = \mathbf{y}_1, \quad \mathbf{u}_1 = \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|} \quad (3.24)$$

$$\mathbf{w}_j = \mathbf{y}_j - \sum_{k=1}^{j-1} \mathbf{u}_k \cdot (\mathbf{u}_k^* \mathbf{y}_j), \quad \mathbf{u}_j = \frac{\mathbf{w}_j}{\|\mathbf{w}_j\|} \quad j = 2, \dots, n.$$

Observe que $\mathbf{u}_i^* \mathbf{u}_j = (\mathbf{u}_j^* \mathbf{u}_i)^*$, pero

$$\mathbf{u}_i^* \mathbf{u}_j = \frac{\mathbf{w}_i^*}{\|\mathbf{w}_i\|} \mathbf{u}_j,$$

y por otra parte

$$(\mathbf{u}_j^* \mathbf{u}_i)^* = \left(\frac{\mathbf{w}_j^*}{\|\mathbf{w}_j\|} \mathbf{u}_i \right)^* = \mathbf{u}_i^* \frac{\mathbf{w}_j}{\|\mathbf{w}_j\|},$$

consecuentemente $\mathbf{u}_i^* \mathbf{u}_j = \delta_{ij}$ para cualesquiera $1 \leq i, j \leq n$, por lo que la matriz $\mathcal{U} = (\mathbf{u}_1 \dots, \mathbf{u}_n)$ pertenece al conjunto $Sp(n)$.

Ahora sea $\mathbf{S} \in Sp(n)$ y considere $\tilde{\mathbf{Y}} = \mathbf{S}\mathbf{Y}$ así como sus respectivos $\tilde{y}_i, \tilde{\mathbf{w}}_i, \tilde{\mathbf{u}}_i$ y $\tilde{\mathbf{U}}$ correspondientes al proceso de ortonormalización de Gram-Schmidt. Se probará por inducción que

$$\tilde{\mathbf{u}}_i = \mathbf{S}\mathbf{u}_i,$$

para $i = 1, \dots, n$.

Primero note que

$$\tilde{\mathbf{u}}_1 = \mathbf{S}\mathbf{u}_1,$$

ahora suponga que dado $2 \leq i \leq n$, se cumple que $\tilde{\mathbf{u}}_j = \mathbf{S}\mathbf{u}_j$ para todo $2 \leq j \leq i-1$ y veamos que $\tilde{\mathbf{u}}_i = \mathbf{S}\mathbf{u}_i$. En efecto, usando la hipótesis de inducción y el hecho de que $\mathbf{S}^* \mathbf{S} = \mathbf{S}\mathbf{S}^* = \mathbf{I}$ se sigue que

$$\begin{aligned} \tilde{\mathbf{w}}_i &= \tilde{\mathbf{y}}_i - \sum_{k=1}^{i-1} \tilde{\mathbf{u}}_k \cdot (\tilde{\mathbf{u}}_k^* \tilde{\mathbf{y}}_i) \\ &= \mathbf{S}\mathbf{y}_i - \sum_{k=1}^{i-1} (\mathbf{S}\mathbf{u}_k) \cdot ((\mathbf{S}\mathbf{u}_k)^* \mathbf{S}\mathbf{y}_i) \\ &= \mathbf{S}\mathbf{w}_i, \end{aligned}$$

con lo que $\tilde{\mathbf{u}}_i = \mathbf{S}\mathbf{u}_i$ y la inducción está completa.

De lo anterior se sigue que $\tilde{\mathbf{U}} = \mathbf{S}\mathbf{U}$, para cualquier $\mathbf{S} \in \mathcal{S}p(n)$. Debido a que las entradas de \mathbf{Y} son i.i.d., la invarianza de \mathbf{S} nos garantiza que $\mathbf{S}\mathbf{Y} \stackrel{\mathcal{L}}{=} \mathbf{Y}$ (véase el Teorema 1 de [48]). Consecuentemente $\mathbf{U} \stackrel{\mathcal{L}}{=} \mathbf{S}\mathbf{U}$, para cualquier $\mathbf{S} \in \mathcal{S}p(n)$, de lo cual se sigue el resultado. ■

Lema 3.10. Sean $\mathbf{z}_k, \mathbf{u}_k, \mathbf{w}_k, \Sigma_k$ y a_k como en (3.33)-(3.36). Para cualquier entero m y reales r, s, t tales que $1 \leq m \leq n, r > 0, s > 0$ y $t > 0$, se cumple la siguiente desigualdad

$$\begin{aligned} \mathbb{P}(\max_{1 \leq k \leq m} \|\sqrt{n}\mathbf{u}_k - \mathbf{z}_k\| \geq rs + t) &\leq \mathbb{P}(\max_{1 \leq k \leq m} |a_k - 1| \geq r) \\ &+ \mathbb{P}(\max_{1 \leq k \leq m} \|\mathbf{w}_k\| \geq s) + \mathbb{P}(\max_{1 \leq k \leq m} \|\Sigma_{k-1}\mathbf{z}_k\| \geq t). \end{aligned}$$

Demostración Note que basta mostrar que

$$\|\sqrt{n}\mathbf{u}_k - \mathbf{z}_k\| \leq |a_k - 1|\|\mathbf{w}_k\| + \|\Sigma_{k-1}\mathbf{z}_k\|, \quad (3.25)$$

se cumple para todo $k = 1, \dots, n$.

En efecto, note que $\mathbf{w}_k = (\mathbf{I} - \Sigma_{k-1})\mathbf{z}_k$, luego

$$\begin{aligned} \sqrt{n}\mathbf{u}_k - \mathbf{z}_k &= a_k\mathbf{w}_k - \mathbf{z}_k = a_k\mathbf{w}_k - \mathbf{w}_k + \mathbf{w}_k - \mathbf{z}_k \\ &= (a_k - 1)\mathbf{w}_k + (\mathbf{w}_k - \mathbf{z}_k) \\ &= (a_k - 1)\mathbf{w}_k - \Sigma_{k-1}\mathbf{z}_k, \end{aligned}$$

de lo cual se sigue la desigualdad que se quería. ■

Lema 3.11. Sea k un entero entre 1 y n , se cumple que

$$\mathbf{A}(\Sigma_{k-1}\mathbf{z}_k) \stackrel{\mathcal{L}}{=} \Sigma_{k-1}\mathbf{z}_k$$

y

$$\mathbf{A}\mathbf{w}_k \stackrel{\mathcal{L}}{=} \mathbf{w}_k,$$

para cualquier $\mathbf{A} \in \mathcal{G}_\beta(n)$ para $\beta = 1, 2$ y 4 .

Demostración Por la invarianza de la matriz $\Gamma = (\mathbf{u}_1, \dots, \mathbf{u}_n)$ para cualquier $\mathbf{A} \in \mathcal{O}(n), \mathcal{U}(n)$ o $\mathcal{S}p(n)$, en cada caso, se tiene que

$$\mathbf{A}(\mathbf{u}_1, \dots, \mathbf{u}_n) = \mathbf{A}\Gamma \stackrel{\mathcal{L}}{=} \Gamma = (\mathbf{u}_1, \dots, \mathbf{u}_n),$$

entonces

$$\mathbf{A}(\mathbf{u}_1, \dots, \mathbf{u}_k) \stackrel{\mathcal{L}}{=} (\mathbf{u}_1, \dots, \mathbf{u}_k),$$

para cualquier $k = 1, \dots, n$.

Lo anterior implica que

$$\begin{aligned} \mathbf{A}(\Sigma_{k-1} \mathbf{z}_k) &= \mathbf{A} \{ (\mathbf{u}_1, \dots, \mathbf{u}_{k-1}) (\mathbf{u}_1, \dots, \mathbf{u}_{k-1})^* \} \mathbf{z}_k \\ &\stackrel{\mathcal{L}}{=} \{ \mathbf{A}(\mathbf{u}_1, \dots, \mathbf{u}_{k-1}) (\mathbf{A}(\mathbf{u}_1, \dots, \mathbf{u}_{k-1}))^* \} \mathbf{A} \mathbf{z}_k \end{aligned}$$

luego

$$\begin{aligned} \mathbf{A} \mathbf{w}_{k-1} &= \mathbf{A}(\mathbf{I} - \Sigma_{k-1}) \mathbf{z}_k \\ &= \mathbf{A} \mathbf{z}_k - \mathbf{A}(\Sigma_{k-1} \mathbf{z}_k) \\ &\stackrel{\mathcal{L}}{=} \mathbf{A} \mathbf{z}_k - \{ \mathbf{A}(\mathbf{u}_1, \dots, \mathbf{u}_{k-1}) (\mathbf{A}(\mathbf{u}_1, \dots, \mathbf{u}_{k-1}))^* \} \mathbf{A} \mathbf{z}_k \\ &= (\mathbf{I} - \{ \mathbf{A}(\mathbf{u}_1, \dots, \mathbf{u}_{k-1}) (\mathbf{A}(\mathbf{u}_1, \dots, \mathbf{u}_{k-1}))^* \}) \mathbf{A} \mathbf{z}_k. \end{aligned}$$

Note que por definición $(\mathbf{u}_1, \dots, \mathbf{u}_{k-1})$ es función de $\mathbf{z}_1, \dots, \mathbf{z}_{k-1}$, por lo que \mathbf{z}_k es independiente de $(\mathbf{u}_1, \dots, \mathbf{u}_{k-1})$, de lo cual es inmediato que

$$\mathbf{A} \mathbf{z}_k \perp \mathbf{A}(\mathbf{u}_1, \dots, \mathbf{u}_{k-1}).$$

Por otra parte, como \mathbf{z}_k tiene entradas independientes y con distribución gaussiana es claro que $\mathbf{A} \mathbf{z}_k \stackrel{\mathcal{L}}{=} \mathbf{z}_k$ entonces,

$$\begin{aligned} \mathbf{A}(\Sigma_{k-1} \mathbf{z}_k) &\stackrel{\mathcal{L}}{=} \{ \mathbf{A}(\mathbf{u}_1, \dots, \mathbf{u}_{k-1}) (\mathbf{A}(\mathbf{u}_1, \dots, \mathbf{u}_{k-1}))^* \} \mathbf{A} \mathbf{z}_k \\ &\stackrel{\mathcal{L}}{=} \{ \mathbf{A}(\mathbf{u}_1, \dots, \mathbf{u}_{k-1}) (\mathbf{A}(\mathbf{u}_1, \dots, \mathbf{u}_{k-1}))^* \} \mathbf{z}_k \\ &\stackrel{\mathcal{L}}{=} \{ (\mathbf{u}_1, \dots, \mathbf{u}_{k-1}) (\mathbf{u}_1, \dots, \mathbf{u}_{k-1})^* \} \mathbf{z}_k \\ &= \Sigma_{k-1} \mathbf{z}_k. \end{aligned}$$

Análogamente se obtiene $\mathbf{A} \mathbf{w}_k \stackrel{\mathcal{L}}{=} \mathbf{w}_k$.

■

Lema 3.12. *Sea \mathbf{A} una matriz de tamaño $n \times n$ con entradas en \mathbb{F}_β , independiente de \mathbf{z}_1 que satisface $\mathbf{A}^* = \mathbf{A}$, $\mathbf{A}^2 = \mathbf{A}$ y $\text{tr}(\mathbf{A}) = k$ para alguna constante $k \in [1, n]$. Si $\mathbf{H}(\mathbf{A} \mathbf{z}_1) \stackrel{\mathcal{L}}{=} \mathbf{A} \mathbf{z}_1$ para alguna $\mathbf{H} \in \mathcal{O}(n), \mathcal{U}(n)$ o $\text{Sp}(n)$, entonces cada entrada de $\mathbf{A} \mathbf{z}_1$ sigue la misma distribución que*

$$\eta \left[\frac{\left(\sum_{i=1}^{\beta k} \xi_i^2 \right)}{\left(\sum_{i=1}^{\beta n} \xi_i^2 \right)} \right]^{\frac{1}{2}},$$

donde η es una variable aleatoria con distribución $\mathbb{FN}(0, 1)$ y las variables $\{\xi_i\}_{i \geq 1}$ son independientes y con distribución gaussiana estándar.

Demostración Si $k = n$, entonces $\mathbf{A} = \mathbf{I}$ y la conclusión resulta inmediata, por lo tanto se supondrá $1 \leq k < n$.

Ahora bien, se sabe que existe $\mathbf{U} \in \mathcal{G}_\beta(n)$ para $\beta = 1, 2$ y 4 respectivamente, tal que

$$\mathbf{A} = \mathbf{U}^* \text{diag}(\mathbf{I}_k, \mathbf{0}) \mathbf{U},$$

de donde \mathbf{U} es función de \mathbf{A} y por tanto $\mathbf{U} \perp \mathbf{z}_1$. Más aún, se probará que \mathbf{Uz}_1 es independiente de \mathbf{z}_1 .

Sean $f(\cdot)$ y $g(\cdot)$ funciones real valuadas, acotadas y Borelianas definidas sobre el espacio de matrices de tamaño $n \times n$ y \mathbb{F}_β^n , respectivamente, entonces

$$\begin{aligned} E[f(\mathbf{U})g(\mathbf{Uz}_1)] &= E\{f(\mathbf{U})E[g(\mathbf{Uz}_1)|\mathbf{U}]\} \\ &= E\{f(\mathbf{U})E[g(\mathbf{Uz}_1)]\} \\ &= E[f(\mathbf{U})]E[g(\mathbf{Uz}_1)], \end{aligned}$$

la última igualdad implica la independencia de \mathbf{U} y \mathbf{Uz}_1 .

Por otra parte, sea $\mathbf{H} \in \mathcal{G}_\beta(n)$ para $\beta = 1, 2, 4$ con distribución de Haar e independiente tanto de \mathbf{U} como de \mathbf{z}_1 tal que

$$\mathbf{H}(\mathbf{Az}_1) \stackrel{\mathcal{L}}{=} \mathbf{Az}_1$$

ya se ha probado que $\mathbf{U} \perp \mathbf{Uz}_1$, por lo tanto $\mathbf{OU}^* \perp \mathbf{Uz}_1$ y

$$\begin{aligned} \mathbf{Az}_1 &\stackrel{\mathcal{L}}{=} \mathbf{HAz}_1 = (\mathbf{HU}^*) \begin{pmatrix} \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} (\mathbf{Uz}_1) \\ &\stackrel{\mathcal{L}}{=} \mathbf{H} \begin{pmatrix} \mathbf{I}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{z}_1 = \mathbf{H}_{n \times k} \begin{pmatrix} z_1 \\ \vdots \\ z_k \end{pmatrix}, \end{aligned}$$

donde $\mathbf{H}_{n \times k}$ son las primeras k columnas de \mathbf{H} y $(z_1, \dots, z_k)^t$ las primeras k entradas de \mathbf{z}_1 . Lo cual nos dice que la primera entrada de \mathbf{Az}_1 tiene la misma distribución que $\sum_{i=1}^k h_{1i} z_i$.

Por la independencia de \mathbf{H} y \mathbf{z}_1 y el Lema 3.9 se sigue que \mathbf{Az}_1 tiene la misma distribución que

$$\frac{\sum_{i=1}^k \eta_i z_i}{\left(\sum_{i=1}^n |\eta_i|^2\right)^{\frac{1}{2}}} = \left(\frac{\sum_{i=1}^k |\eta_i|^2}{\sum_{i=1}^n |\eta_i|^2}\right)^{\frac{1}{2}} \sum_{i=1}^k \lambda_i z_i,$$

donde $\{\eta_i\}_{i=1}^n$ son variables aleatorias independientes y con distribución $\mathbb{FN}(0, 1)$ y $\lambda_i = \eta_i \left(\sum_{i=1}^k |\eta_i|^2 \right)^{-\frac{1}{2}}$.

Es claro que $\sum_{i=1}^k \lambda_i z_i$ tiene distribución $\mathbb{FN}(0, 1)$ y por la independencia de las variables η_i se sigue que $|\eta_i|^2$ se distribuye como $\frac{\chi^2(\beta)}{\beta}$ para $i = 1, \dots, n$.

Para terminar la prueba resta verificar que $\sum_{i=1}^k \lambda_i z_i$ es independiente de $\frac{\sum_{i=1}^k |\eta_i|^2}{\sum_{i=1}^n |\eta_i|^2}$, note que las variables z_i, η_i para $i = 1, \dots, n$ son independientes e idénticamente distribuidas y escribiendo la suma por partes como,

$$\sum_{i=1}^n |\eta_i|^2 = \sum_{i=1}^k |\eta_i|^2 + \sum_{i=k+1}^n |\eta_i|^2,$$

por lo que es suficiente probar la independencia de $\sum_{i=1}^k \lambda_i z_i$ y $\sum_{i=1}^k |\eta_i|^2$.

Sean $f(\cdot)$ y $g(\cdot)$, funciones Borelianas y acotadas, definidas sobre \mathbb{F}_β y \mathbb{R} respectivamente, entonces

$$E \left[f \left(\sum_{i=1}^k \lambda_i z_i \right) g \left(\sum_{i=1}^k |\eta_i|^2 \right) \right] = E \left[f \left(\sum_{i=1}^k \lambda_i z_i \right) \right] E \left[g \left(\sum_{i=1}^k |\eta_i|^2 \right) \right]$$

lo cual implica la independencia deseada. ■

Lema 3.13. *Siguiendo la notación del lema anterior, se cumplen las siguientes afirmaciones*

i) *Cada entrada del vector columna $\Sigma_{k-1} \mathbf{z}_k$ tiene la misma distribución que*

$$\eta \left(\frac{\sum_{i=1}^{\beta(k-1)} \xi_i^2}{\sum_{i=1}^{\beta n} \xi_i^2} \right)^{\frac{1}{2}},$$

para $k = 2, \dots, n$.

ii) *Cada entrada del vector columna \mathbf{w}_k tiene la misma distribución que*

$$\eta \left(\frac{\sum_{i=1}^{\beta(n-k+1)} \xi_i^2}{\sum_{i=1}^{\beta n} \xi_i^2} \right)^{\frac{1}{2}},$$

para $k = 2, \dots, n$.

iii) *Se cumple que*

$$\|\mathbf{w}_k\|^2 \stackrel{\mathcal{L}}{=} \beta^{-1} \chi^2(\beta(n-k+1)),$$

para toda $k = 1, \dots, n$.

Demostración Recuerde que

$$\Sigma_k = (\mathbf{u}_1, \dots, \mathbf{u}_{k-1})(\mathbf{u}_1, \dots, \mathbf{u}_{k-1})^*$$

para $k = 2, \dots, n$, y

$$\mathbf{w}_k = (\mathbf{I} - \Sigma_{k-1})\mathbf{z}_k$$

para $k = 1, \dots, n$.

Directamente de la definición de Σ_{k-1} es fácil ver que tanto Σ_{k-1} como $(\mathbf{I} - \Sigma_{k-1})$ son matrices unitarias e idempotentes. ASÍ que para que satisfacen las hipótesis del lema anterior, sólo falta probar que la traza de estas matrices es igual a algún entero.

En efecto

$$\text{tr}(\Sigma_{k-1}) = \text{tr} \left(\sum_{i=1}^{k-1} \mathbf{u}_i \mathbf{u}_i^* \right) = \sum_{i=1}^{k-1} \text{tr}(\mathbf{u}_i \mathbf{u}_i^*) = k - 1,$$

y entonces

$$\text{tr}(\mathbf{I} - \Sigma_{k-1}) = \text{tr}(\mathbf{I}) - \text{tr}(\Sigma_{k-1}) = n - k + 1,$$

con lo cual el resultado se sigue del lema anterior. ■

Lema 3.14. Sea \mathbb{F}_β , con $\beta = 1, 2$ y 4 , entonces

$$\mathbb{P} \left(\max_{1 \leq i \leq m} |a_i - 1| \geq r \right) \leq 4me^{-\frac{\beta nr^2}{16}},$$

para toda $r \in (0, \frac{1}{4})$ y $m \leq \frac{nr}{2}$.

Demostración Hay que notar que

$$\mathbb{P} \left(\max_{1 \leq i \leq m} |a_i - 1| \geq r \right) \leq m \max_{1 \leq i \leq m} \mathbb{P}(|a_i - 1| \geq r).$$

Para toda $i = 1, \dots, n$ se tiene que

$$\begin{aligned} \mathbb{P}(|a_i - 1| \geq r) &\leq \mathbb{P}(a_i - 1 \geq r) + \mathbb{P}(a_i - 1 \leq -r) \\ &= \mathbb{P} \left(\frac{\sqrt{n}}{\|\mathbf{w}_i\|} \geq 1 + r \right) + \mathbb{P} \left(\frac{\sqrt{n}}{\|\mathbf{w}_i\|} \leq 1 - r \right) \\ &= \mathbb{P} \left(\frac{\|\mathbf{w}_i\|^2}{n} \leq (1 + r)^{-2} \right) + \mathbb{P} \left(\frac{\|\mathbf{w}_i\|^2}{n} \leq (1 - r)^{-2} \right) \\ &\leq \mathbb{P} \left(\frac{\|\mathbf{w}_i\|^2}{n} \leq 1 - r \right) + \mathbb{P} \left(\frac{\|\mathbf{w}_i\|^2}{n} \leq 1 + 2r \right) \end{aligned} \tag{3.26}$$

donde la última desigualdad se sigue del inciso (iii) de C.4 y se cumple siempre que $r \in (0, \frac{1}{4})$. Ahora basta acotar las dos últimas probabilidades.

De (iii) del Lema 3.13 se sigue que

$$\frac{\|\mathbf{w}_i\|^2}{n} \sim \frac{\sum_{i=1}^{\beta(n-i+1)} \xi_i^2}{n\beta} \geq \frac{\sum_{i=1}^{\beta(n-m)} \xi_i^2}{n\beta}$$

para toda $i = 1, \dots, m$, entonces

$$\max_{1 \leq i \leq m} \mathbb{P} \left(\frac{\|\mathbf{w}_i\|^2}{n} \leq 1 - r \right) \leq \mathbb{P} \left(\frac{\sum_{i=1}^{\beta(n-m)} \xi_i^2}{(n-m)\beta} \leq \frac{n(1-r)}{(n-m)} \right).$$

Haciendo $b = \frac{n(1-r)}{(n-m)}$ y $B = (-\infty, b]$, del Lema B.1, observando que $b \in (0, 1)$ siempre que $m \leq \frac{nr}{2}$ y usando el hecho de que $I(x)$ es decreciente en el intervalo $(0, 1)$ puede concluirse que

$$\begin{aligned} \max_{1 \leq i \leq m} \mathbb{P} \left(\frac{\|\mathbf{w}_i\|^2}{n} \leq 1 - r \right) &\leq \mathbb{P} \left(\frac{\sum_{i=1}^{\beta(n-m)} \xi_i^2}{(n-m)\beta} \leq \frac{n(1-r)}{(n-m)} \right) \\ &\leq 2e^{-\beta(n-m)I(B)} \\ &= 2e^{-\beta(n-m)I(b)}. \end{aligned} \tag{3.27}$$

para $r \in (0, \frac{1}{4})$ y $m \leq \frac{nr}{2}$.

Ahora bien, del inciso (i) de Lema C.4,

$$(n-m)I(B) \geq (n-m) \cdot \frac{(1-b)^2}{4} \geq \frac{(nr-m)^2}{4(n-m)} \geq \frac{nr^2}{16} \tag{3.28}$$

siempre que $1 \leq m \leq \frac{nr}{2}$.

Así, de (3.27) y (3.28) se tiene la siguiente desigualdad

$$\max_{1 \leq i \leq m} \mathbb{P} \left(\frac{\|\mathbf{w}_i\|^2}{n} \leq 1 - r \right) \leq 2e^{-\frac{\beta nr^2}{16}}. \tag{3.29}$$

Por otra parte, del inciso (iii) del Lema 3.13 y considerando

$$\lambda := \inf_{x \geq 1+2r} I(x)$$

con $I(x)$ como en el Lema B.1, entonces se sigue que

$$\|\mathbf{w}\|^2 \sim \frac{\sum_{i=1}^{\beta(n-i+1)} \xi_i^2}{\beta} \leq \frac{\sum_{i=1}^{\beta n} \xi_i^2}{\beta}$$

para todo $i = 1, \dots, m$.

Por lo tanto, para toda $r \in (0, \frac{1}{4})$

$$\begin{aligned} \max_{1 \leq i \leq m} \mathbb{P} \left(\frac{\|\mathbf{w}_i\|^2}{n} \geq 1 + 2r \right) &\leq \mathbb{P} \left(\frac{\sum_{i=1}^{\beta n} \xi_i^2}{n\beta} \geq 1 + 2r \right) \\ &\leq 2e^{-\beta n \lambda}. \end{aligned}$$

Además note que $I(x)$ es creciente en $[1, \infty)$, luego

$$\lambda = I(1 + 2r) = \frac{2r - \log(1 + 2r)}{2} \geq \frac{r^2}{2} \geq \frac{r^2}{16}$$

para todo $r \in (0, \frac{1}{4})$, por lo tanto puede concluirse que

$$\max_{1 \leq i \leq m} \mathbb{P} \left(\frac{\|\mathbf{w}_i\|^2}{n} \geq 1 + 2r \right) \leq 2e^{-\frac{\beta n r^2}{16}} \quad r \in (0, \frac{1}{4}). \quad (3.30)$$

Finalmente, de las ecuaciones (3.26) (3.29) y (3.30) se concluye que

$$\mathbb{P}(|a_i - 1| \geq r) \leq 4e^{-\frac{\beta n r^2}{16}}$$

siempre que $r \in (0, \frac{1}{4})$ y $m \leq \frac{nr}{2}$.

■

Lema 3.15. Sea \mathbb{F}_β , con $\beta = 1, 2$ y 4 , $m = 1, \dots, n$ y $t > 0$, se cumple que

$$\mathbb{P} \left(\max_{1 \leq i \leq m} \|\Sigma_{i-1} \mathbf{z}_i\| \geq t \right) \leq 3mn \chi_\beta(n, t) \left(1 + \frac{t^2}{3(m + t\sqrt{n})} \right)^{-\frac{\beta n}{2}},$$

donde

$$\chi_\beta(n, t) = \begin{cases} t^{-1}, & \text{si } \beta = 1 \\ 1, & \text{si } \beta = 2 \\ 8nt^2 + 1, & \text{si } \beta = 4. \end{cases}$$

Demostración Recuerde que $\Sigma_0 = 0$, por lo que se puede suponer sin pérdida de generalidad que $1 < m < n$.

Del inciso (i) de 3.13 se tiene que

$$\begin{aligned} \mathbb{P} \left(\max_{1 \leq i \leq m} \|\Sigma_{i-1} \mathbf{z}_i\| \geq t \right) &\leq mn \max_{2 \leq i \leq m} \mathbb{P}(|\eta| W_i \geq t) \\ &\leq mn \mathbb{P}(|\eta| W_{m+1} \geq t), \end{aligned} \quad (3.31)$$

donde $\eta \sim \mathbb{NF}_2(0, 1)$ y las variables

$$W_i = \left(\frac{\sum_{k=1}^{\beta(i-1)} \xi_k^2}{\sum_{k=1}^{\beta n} \xi_k^2} \right)^{\frac{1}{2}} \leq 1, \quad i = 2, \dots, n+1$$

son independientes.

Del Lema B.2

$$P(|\mathbb{FN}_\beta(0, 1)| \geq x) \leq \varphi_\beta(x) e^{-\frac{\beta x^2}{2}},$$

donde

$$\varphi_\beta(x) = \begin{cases} \frac{1}{x}, & \text{si } \beta = 1 \\ 1, & \text{si } \beta = 2 \\ 2x^2 + 1, & \text{si } \beta = 4. \end{cases},$$

esto es

$$P(|\eta| \cdot W_{m+1} \geq t) \leq E \left\{ \varphi_\beta(tW_{m+1}^{-1} \exp\left(-\frac{\beta t^2}{2} W_{m+1}^2\right)) \right\}. \quad (3.32)$$

Por lo tanto se tienen los siguientes casos:

i) Si $\beta = 1$, entonces

$$\varphi_\beta(tW_{m+1}^{-1}) \leq \frac{1}{t},$$

así que considerando $a = \frac{\beta t^2}{2}$ y reemplazando m por βm así como n por βn en el Lema B.3 se obtiene

$$P(|\eta| \cdot W_{m+1} \geq t) \leq \frac{3}{t} \left(1 + \frac{\beta t^2}{3(\beta m + t\beta\sqrt{n})}\right)^{-\frac{\beta n}{2}} \leq \frac{3}{t} \left(1 + \frac{t^2}{3(\beta m + t\sqrt{n})}\right)^{-\frac{\beta n}{2}}.$$

ii) Si $\beta = 2$, entonces

$$\varphi_\beta(tW_{m+1}^{-1}) = 1,$$

así que considerando $a = \frac{\beta t^2}{2}$ y reemplazando m por βm así como n por βn en el Lema B.3 se obtiene

$$P(|\eta| \cdot W_{m+1} \geq t) \leq E \exp\left\{-\frac{\beta t^2}{2} W_{m+1}^{-2}\right\} \leq 3 \left(1 + \frac{t^2}{3(m + t\sqrt{n})}\right)^{-\frac{\beta n}{2}}.$$

iii) Si $\beta = 4$, entonces

$$\varphi_\beta(x) = 2x^2 + 1,$$

además note que $\varphi_\beta(x)$ es creciente, mientras que $\psi(x) := e^{-\frac{\beta x^2}{2}}$ es decreciente en $[0, \infty)$. Del Lema B.3 se obtiene

$$\begin{aligned} P(|\eta| \cdot W_{m+1} \geq t) &\leq E \left\{ \varphi_\beta(tW_{m+1}^{-1}) \right\} E \exp\left\{-\frac{\beta t^2}{2} W_{m+1}^{-2}\right\} \\ &\leq 3(8nt^2 + 1) \left(1 + \frac{t^2}{3(m + t\sqrt{n})}\right)^{-\frac{\beta n}{2}}. \end{aligned}$$

■

3.3.3. Prueba del Resultado Principal

Primero considere una matriz aleatoria $\mathbf{Z}_n = (z_{ij})$ en $\text{GinOE}(n)$, $\text{GinUE}(n)$ o $\text{GinSE}(n)$, escribimos

$$\mathbf{Z}_n = (\mathbf{z}_1, \mathbf{z}_2, \dots, \mathbf{z}_n), \quad (3.33)$$

es decir $\{\mathbf{z}_r\}_{1 \leq r \leq n}$ son los vectores columna de \mathbf{Z}_n .

Ahora bien, sea Γ_n la matriz formada por los vectores columna resultantes de aplicar el proceso de ortonormalización de Gram-Schmidt a los vectores columna de \mathbf{Z}_n . Siguiendo la notación del Capítulo 2 tenemos

$$\Gamma_n = (\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n), \quad (3.34)$$

donde los vectores columna \mathbf{u}_j son como en (2.1).

De esta forma tenemos que \mathbf{Z}_n tiene entradas gaussianas $\mathbb{N}\mathbb{F}_\beta(0, 1)$ y además se sabe que la matriz Γ_n tiene distribución de Haar sobre $\mathcal{G}_\beta(n)$ (Véase Cáp. 2) con lo que Γ_n y \mathbf{Z}_n satisfacen las condiciones (i) y (ii) del Teorema 3.8.

Para continuar con la demostración hay que considerar las cantidades

$$\Sigma_0 = 0, \quad \Sigma_k = (\mathbf{u}_1, \dots, \mathbf{u}_k)(\mathbf{u}_1, \dots, \mathbf{u}_k)^*, \quad k = 1, \dots, n \quad (3.35)$$

$$a_k = \frac{\sqrt{n}}{\|\mathbf{w}_k\|}, \quad (3.36)$$

entonces se tiene que

$$\begin{aligned} \sum_{i=1}^{k-1} \mathbf{u}_i(\mathbf{u}_i^* \mathbf{z}_k) &= \{(\mathbf{u}_1, \dots, \mathbf{u}_k)(\mathbf{u}_1, \dots, \mathbf{u}_k)^*\} \mathbf{z}_k \\ &= \Sigma_{k-1} \mathbf{z}_k \end{aligned}$$

para $k = 2, \dots, n$, luego

$$\mathbf{w}_k = \mathbf{z}_k - \sum_{i=1}^{k-1} \mathbf{u}_i(\mathbf{u}_i^* \mathbf{z}_k) = (\mathbf{I} - \Sigma_{k-1}) \mathbf{z}_k. \quad (3.37)$$

Estas expresiones serán de gran utilidad en lo sucesivo.

Prueba del Teorema 3.8

Demostración Primero note que del Lema 3.10 se tiene que,

$$\begin{aligned} \mathbb{P}\left(\max_{1 \leq k \leq m} \|\sqrt{n} \mathbf{u}_k - \mathbf{z}_k\| \geq rs + t\right) &\leq \mathbb{P}\left(\max_{1 \leq k \leq m} |a_k - 1| \geq r\right) \\ &\quad + \mathbb{P}\left(\max_{1 \leq k \leq m} \|\mathbf{w}_k\| \geq s\right) + \mathbb{P}\left(\max_{1 \leq k \leq m} \|\Sigma_{k-1} \mathbf{z}_k\| \geq t\right) \end{aligned}$$

para $r > 0, s > 0$ y $t > 0$.

Para acotar el segundo término, note que del Lema 3.13 para todo $k = 1, \dots, n$,

$$\mathbf{w}_k \stackrel{\mathcal{L}}{=} \eta \left(\frac{\sum_{i=1}^{\beta(n-k+1)} \xi_i^2}{\sum_{i=1}^{\beta n} \xi_i^2} \right)^{\frac{1}{2}},$$

y además se satisface la siguiente desigualdad

$$\frac{\sum_{i=1}^{\beta(n-k+1)} \xi_i^2}{\sum_{i=1}^{\beta n} \xi_i^2} \leq 1.$$

Por lo tanto

$$\mathbb{P}(\max_{1 \leq k \leq m} \|\mathbf{w}_k\| \geq s) \leq m \max_{1 \leq k \leq m} \mathbb{P}(\|\mathbf{w}_k\| \geq s).$$

De la desigualdad del triángulo y el Lema B.2 se obtiene

$$\begin{aligned} \mathbb{P}(\max_{1 \leq k \leq m} \|\mathbf{w}_k\| \geq s) &\leq m \max_{1 \leq k \leq m} \mathbb{P}(\|\mathbf{w}_k\| \geq s) \\ &\leq mn \mathbb{P}(\|\mathbb{FN}(0, 1)\| \geq s) \\ &\leq mn \varphi_\beta(s) e^{-\frac{\beta s^2}{2}} \end{aligned}$$

donde

$$\varphi_\beta(x) = \begin{cases} \frac{1}{x}, & \text{si } \beta = 1; \\ 1, & \text{si } \beta = 2; \\ 2x^2 + 1, & \text{si } \beta = 4. \end{cases}$$

Por otra parte, note que por los Lemas 3.15 y 3.14 se tienen cotas para el primer y tercer término del lado izquierdo respectivamente, por lo que se concluye que

$$\begin{aligned} &\mathbb{P}(\max_{1 \leq k \leq m} \|\sqrt{n}\mathbf{u}_k - \mathbf{z}_k\| \geq rs + t) \\ &\leq 4me^{-\frac{\beta nr^2}{16}} + mn \varphi_\beta(s) e^{-\frac{\beta s^2}{2}} + 3mn \chi_\beta \left(1 + \frac{1}{3(m + t\sqrt{n})} \right)^{-\frac{\beta n}{2}}, \end{aligned}$$

para todo $r \in (0, \frac{1}{4})$, $s > 0$, $t > 0$ y $1 \leq m \leq \frac{nr}{2}$.

Para obtener el resultado deseado, basta notar que

$$\varphi_\beta(s) \leq \chi_\beta(n, s)$$

para toda $s > 0$, $n \geq 1$ y $\beta = 1, 2, 4$.

■

Capítulo 4

Consecuencias sobre la Distribución Espectral Asintótica de Bloques

El objetivo de este capítulo es mostrar algunas consecuencias de los resultados que fueron expuestos en el capítulo anterior. De acuerdo a los dos paradigmas de aproximación estudiados la distribución conjunta de matrices aleatorias de Haar truncadas de dimensión suficientemente grande es cercana a la de matrices con entradas independientes y gaussianas estándar. Una de las razones por las que este hecho tiene relevancia es que permitirá extender resultados que se cumplen para matrices gaussianas de dimensión grande a matrices con distribución de Haar.

Comenzaremos con el resultado siguiente, consecuencia directa de la aproximación en el sentido de distancia variacional probada en los Teoremas 3.2 y 3.3 y será la clave para obtener algunas propiedades asintóticas sobre la distribución espectral de matrices aleatorias de Haar truncadas.

Corolario 2. *Bajo las condiciones del Teorema 3.2, considere $\{f_n\}_{n \leq 1}$ una sucesión de funciones medibles, $f_n(\cdot) = f(x_1, \dots, x_n) : \mathbb{R}^n \rightarrow \mathbb{R}^k$. Si existe $D \in \mathcal{B}(\mathbb{R}^k)$ tal que*

$$\lim_{n \rightarrow \infty} P(f_n(\lambda_1(\mathbf{W}_n), \dots, \lambda_n(\mathbf{W}_n)) \in D) = c,$$

para alguna constante c , donde $\mathbf{W}_n = \mathbf{Z}^ \mathbf{Z}$, entonces también se cumple que*

$$\lim_{n \rightarrow \infty} P(f_n(\lambda_1(\mathcal{W}_n), \dots, \lambda_n(\mathcal{W}_n)) \in D) = c,$$

donde $\mathcal{W}_n = n\mathbf{U}^ \mathbf{U}$.*

Demostración. Sea n entero positivo fijo, del Teorema de Perturbación para valores singulares se tiene que para cualesquiera \mathbf{A}, \mathbf{B} matrices de tamaño $p \times q$ se cumple que

$$\begin{aligned} \max_{1 \leq i \leq q} |\sqrt{\lambda_i(\mathbf{A}^* \mathbf{A})} - \sqrt{\lambda_i(\mathbf{B}^* \mathbf{B})}| &\leq \max_{1 \leq i \leq q} \sqrt{\lambda_i(\mathbf{A} - \mathbf{B})^* (\mathbf{A} - \mathbf{B})} \\ &= \sqrt{\text{tr}((\mathbf{A} - \mathbf{B})^* (\mathbf{A} - \mathbf{B}))} \end{aligned}$$

lo cual implica que $(\lambda_1(\mathbf{A}^* \mathbf{A}), \dots, \lambda_q(\mathbf{A}^* \mathbf{A}))$ es un vector continuo de \mathbf{A} . Por lo tanto, existe $E \in \mathcal{B}(\mathbb{R}^{pq})$ tal que

$$\{f_n(\lambda_1(\mathbf{W}_n), \dots, \lambda_q(\mathbf{W}_n)) \in D\} = \{\mathbf{Z} \in E\}$$

y

$$\{f_n(\lambda_1(\mathcal{W}_n), \dots, \lambda_q(\mathcal{W}_n)) \in D\} = \{\mathbf{U} \in E\}.$$

El Teorema 3.2 garantiza que

$$P(\{\mathbf{Z} \in E\}) - P(\{\mathbf{U} \in E\}) \rightarrow 0$$

cuando $n \rightarrow \infty$, por lo que se sigue el resultado que se quería. ■

Los resultados siguientes son válidos para matrices aleatorias con distribución de Haar sobre $\mathcal{O}(n)$ o $\mathcal{U}(n)$, sin embargo las demostraciones se hacen sólo para el caso unitario, el caso ortogonal se sigue por analogía. En las subsecciones siguientes se usa la notación del Teorema 3.2.

4.1. Ley de Grandes Números para el Radio Espectral

La convergencia del radio espectral escalado de matrices con entradas independientes y con distribución gaussiana estándar fue probada por primera vez por Geman [27]. Posteriormente, Yin, Bai y Krishnaiah [78] generalizaron el resultado para matrices con entradas i.i.d. y con cuarto momento finito. Por otra parte, los primeros resultados acerca del comportamiento del menor eigenvalor de matrices aleatorias fueron desarrollados por Yin, Bai y Krishnaiah [77] y Silverstein [61] quien estudió la convergencia casi segura del menor eigenvalor de matrices aleatorias de tipo Wishart.

Teorema 4.1. (Radio espectral) Sea Γ_n una matriz con distribución de Haar en $\mathcal{O}(n)$ o $\mathcal{U}(n)$ y \mathbf{U}_n su bloque superior izquierdo de tamaño $p_n \times q_n$. Sea $\lambda_{\max}(\mathbf{U}_n^* \mathbf{U}_n)$ el máximo eigenvalor de $\mathbf{U}_n^* \mathbf{U}_n$, y $\lambda_{\min}(\mathbf{U}_n^* \mathbf{U}_n)$ definido como

$$\lambda_{\min}(\mathbf{U}_n^* \mathbf{U}_n) = \begin{cases} \lambda_1(\mathbf{U}_n^* \mathbf{U}_n), & \text{si } p_n \geq q_n; \\ \lambda_{(q-p+1)}(\mathbf{U}_n^* \mathbf{U}_n), & \text{si } p_n < q_n. \end{cases}$$

Si $p_n \rightarrow \infty$, $p_n = o(\sqrt{n})$ y $\frac{q_n}{p_n} \rightarrow c \in (0, +\infty)$, entonces

$$\left(\frac{n}{p_n}\right) \lambda_{\max}(\mathbf{U}_n^* \mathbf{U}_n) \rightarrow (1 + \sqrt{c})^2$$

y

$$\left(\frac{n}{p_n}\right) \lambda_{\min}(\mathbf{U}_n^* \mathbf{U}_n) \rightarrow (1 - \sqrt{c})^2$$

en probabilidad, cuando $n \rightarrow \infty$.

Demostración. Considere \mathbf{Z}_n como en el Teorema 3.2. Por hipótesis $\frac{q_n}{p_n} \rightarrow c \in (0, +\infty)$, por lo tanto se satisfacen las hipótesis del Teorema 2.16 de [3] por lo que

$$\left(\frac{n}{p}\right) \lambda_{\max}(\mathbf{Z}^* \mathbf{Z}) \rightarrow (1 + \sqrt{c})^2 \quad (4.1)$$

y

$$\left(\frac{n}{p}\right) \lambda_{\min}(\mathbf{Z}^* \mathbf{Z}) \rightarrow (1 - \sqrt{c})^2 \quad (4.2)$$

en probabilidad, cuando $n \rightarrow \infty$.

Por lo tanto

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\left| \frac{\lambda_{\max}(\mathbf{Z}^* \mathbf{Z})}{p} - (1 + \sqrt{c})^2 \right| \geq \varepsilon \right) \rightarrow 0.$$

Considerando la convergencia anterior junto con la sucesión de funciones

$$f_n(\lambda_1(\mathbf{Z}^* \mathbf{Z}), \dots, \lambda_q(\mathbf{Z}^* \mathbf{Z})) = \lambda_{\min}(\mathbf{Z}^* \mathbf{Z}),$$

el Corolario 2 nos proporciona el resultado deseado. ■

4.2. Ley de Tracy-Widom

El resultado siguiente se refiere a la convergencia de la distribución del máximo eigenvalor de matrices aleatorias tipo Wishart formadas a partir de matrices aleatorias de Haar truncadas hacia las leyes de Tracy Widom.

Teorema 4.2. (*Ley Tracy-Widom*) Sea $\mathbf{\Gamma}_n$ una matriz con distribución de Haar invariante con \mathbf{U}_n su bloque superior izquierdo de tamaño $p_n \times q_n$ y $\lambda_{\max}(\mathbf{U}_n^* \mathbf{U}_n)$ el mayor de los eigenvalores de $\mathbf{U}_n^* \mathbf{U}_n$. Si $p_n \rightarrow \infty$, $p_n = o(\sqrt{n})$ y $q_n/p_n \rightarrow c \in (0, \infty)$, entonces se cumplen las siguientes afirmaciones

i) Si $\mathbf{\Gamma}_n \in \mathcal{O}(n)$ entonces, cuando $n \rightarrow \infty$

$$\frac{n\lambda_{\max}(\mathbf{U}_n^* \mathbf{U}_n) - a_n}{b_n} \text{ converge débilmente a la ley de Tracy-Widom } F_1 \text{ definida como en 1.5,}$$

ii) Si $\Gamma_n \in \mathcal{U}(n)$ entonces, cuando $n \rightarrow \infty$

$$\frac{n\lambda_{\max}(\mathbf{U}_n^* \mathbf{U}_n) - a_n}{b_n} \quad \text{converge débilmente a la ley de Tracy-Widom } F_2 \text{ definida como en 1.5.}$$

donde

$$a_n = (\sqrt{p_n} + \sqrt{q_n})^2 \quad y \quad b_n = (\sqrt{p_n} + \sqrt{q_n}) \left(\frac{1}{\sqrt{p_n}} + \frac{1}{\sqrt{q_n}} \right)^{1/3}.$$

Demostración. Sea \mathbf{Z}_n una matriz de tamaño $p_n \times q_n$ con entradas i.i.d. con ley gaussiana compleja estándar.

Hay que recordar que Johansson [41] probó que

$$\frac{n\lambda_{\max}(\mathbf{Z}_n^* \mathbf{Z}_n) - a_n}{b_n} \quad \text{converge débilmente a la ley de Tracy-Widom } F_2,$$

cuando $n \rightarrow \infty$.

Por otra parte, Johnstone [42] mostró que si las entradas de \mathbf{Z}_n tienen ley gaussiana estándar, entonces se cumple que

$$\frac{n\lambda_{\max}(\mathbf{Z}_n^* \mathbf{Z}_n) - a_n}{b_n} \quad \text{converge débilmente a la ley de Tracy-Widom } F_1,$$

cuando $n \rightarrow \infty$.

Por lo que para probar el resultado basta apelar al Corolario 2 considerando

$$f_n(x_1, \dots, x_q) = \frac{\max_{1 \leq i \leq q} x_i - a_n}{b_n} \quad x_1, \dots, x_q \geq 0,$$

y $D = (-\infty, x]$ para cualquier $x \in \mathbb{R}$.

■

Para exponer otras consecuencias sobre la distribución espectral de los bloques se requiere introducir la siguiente métrica.

Definición 11. Sean μ y ν medidas de probabilidad sobre $(\mathbb{C}, \mathcal{B}(\mathbb{C}))$, definimos

$$\eta(\mu, \nu) := \sup \left\{ \left| \int_{\mathbb{C}} f(x) \mu(dx) - \int_{\mathbb{C}} f(x) \nu(dx) \right| : \|f\|_{BL} \leq 1 \right\}, \quad (4.3)$$

donde

$$\|f\|_{BL} := \|f\|_{\infty} + \|f\|_L$$

para funciones f Lipschitz y acotadas sobre \mathbb{C} y donde $\|f\|_\infty$ y $\|f\|_L$ denotan la norma supremo y la seminorma Lipschitz usuales, es decir

$$\|f\|_\infty := \sup_{x \in \mathbb{C}} |f(x)| \quad \text{y} \quad \|f\|_L := \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}.$$

Es fácil ver que esta métrica genera la topología de convergencia débil de medidas de probabilidad sobre \mathbb{C} , por lo que una sucesión μ_n converge débilmente a μ si y sólo si $\eta(\mu_n, \mu)$ converge a cero cuando n tiende a infinito.

Lema 4.3. *Sea \mathbf{X} una matriz de tamaño $n \times n$ con entradas complejas, $\mu_{\mathbf{X}}$ la distribución espectral empírica de sus eigenvalores $\lambda_1, \dots, \lambda_n$ y μ alguna medida de probabilidad dada. Entonces se cumple que*

$$\eta(\mu_{\mathbf{X}}, \mu),$$

definida como en (4.3) es una función continua en las entradas de \mathbf{X} .

Demostración. Primero note que para toda función f Lipschitz y acotada sobre \mathbb{C} se cumple que

$$\int_{\mathbb{R}} f(x) \mu_{\mathbf{X}}(dx) = \frac{1}{n} \sum_{i=1}^n f(\lambda_i(\mathbf{X})).$$

Ahora bien, sean $\mathbf{X} = (x_{ij})_{1 \leq i, j \leq n}$ y $\mathbf{Y} = (y_{ij})_{1 \leq i, j \leq n}$ matrices aleatorias y π alguna permutación de $\{1, \dots, n\}$

$$\begin{aligned} |\eta(\mu_{\mathbf{X}}, \mu) - \eta(\mu_{\mathbf{Y}}, \mu)| &\leq \frac{1}{n} \sup_{\|f\|_{BL} \leq 1} \left| \sum_{i=1}^n f(\lambda_{\pi(i)}(\mathbf{X})) - f(\lambda_i(\mathbf{Y})) \right| \\ &\leq \max_{1 \leq i \leq n} \sup_{\|f\|_{BL} \leq 1} |f(\lambda_{\pi(i)}(\mathbf{X})) - f(\lambda_i(\mathbf{Y}))| \\ &= \max_{1 \leq i \leq n} |\lambda_{\pi(i)}(\mathbf{X}) - \lambda_i(\mathbf{Y})|, \end{aligned}$$

como esto es cierto para cualquier permutación π , entonces

$$|\eta(\mu_{\mathbf{X}}, \mu) - \eta(\mu_{\mathbf{Y}}, \mu)| \leq \min_{\pi} \max_{1 \leq i \leq n} |\lambda_{\pi(i)}(\mathbf{X}) - \lambda_i(\mathbf{Y})|.$$

Por el Lema C.5 (Véase Apéndice C) se tiene que

$$\min_{\pi} \max_{1 \leq i \leq n} |\lambda_{\pi(i)}(\mathbf{X}) - \lambda_i(\mathbf{Y})| \leq 2^{2-1/n} \|\mathbf{X} - \mathbf{Y}\|_2^{1/n} (\|\mathbf{X}\|_2 + \|\mathbf{Y}\|_2)^{1-1/n},$$

donde $\|\cdot\|_2$ denota la norma usual de matrices. De las últimas desigualdades y el hecho de que

$$\|\mathbf{X}\|_2 \leq \sum_{1 \leq i, j \leq n} |x_{ij}|^2,$$

se sigue que

$$\begin{aligned} |\eta(\mu_{\mathbf{X}}, \mu) - \eta(\mu_{\mathbf{Y}}, \mu)| &\leq 2^{2-1/n} \|\mathbf{X} - \mathbf{Y}\|_2^{1/n} (\|\mathbf{X}\|_2 + \|\mathbf{Y}\|_2)^{1-1/n} \\ &= 2 \left(\sum_{1 \leq i, j \leq n} |x_{ij} - y_{ij}|^2 \right)^{1/2n} (\|\mathbf{X}\|_2 + \|\mathbf{Y}\|_2)^{1-1/n} \end{aligned}$$

con lo que se obtiene la continuidad deseada. ■

4.3. Ley del Círculo

Como se mencionó en el Capítulo 1, la Ley del Círculo surgió como una conjetura, que asegura que la distribución espectral (normalizada adecuadamente) de una matriz aleatoria no simétrica converge a la distribución uniforme sobre el disco unitario, y ha sido ampliamente estudiada para el caso de matrices aleatorias con entradas independientes y gaussianas.

En 2000 Zyczkowski y Sommers [81] estudiaron bloques de tamaño $p \times p$ de matrices aleatorias con distribución de Haar sobre los grupos $\mathcal{G}_\beta(n)$ para $\beta = 1, 2$, simularon la distribución empírica de sus eigenvalores y notaron que todos ellos se encontraban en el disco unitario. Además de esto, observaron que las propiedades de las matrices truncadas quedan determinadas por el valor del cociente p/n . Más tarde, en 2004 Petz y Réffy [56] probaron un resultado interesante en esta misma dirección, consideran bloques de tamaño $m \times n$ de matrices aleatorias con distribución de Haar sobre $\mathcal{U}(n)$ tales que $m/n \rightarrow \gamma \in (0, \infty)$ cuando $n \rightarrow \infty$ y demuestran que su distribución empírica espectral converge en probabilidad a una ley que no es la del círculo. Finalmente, en 2009 Jiang [39] prueba el resultado que se enuncia a continuación, donde se asegura que las condiciones $p_n \rightarrow \infty$ y $p_n = o(\sqrt{n})$ son suficientes para garantizar que la distribución empírica de los eigenvalores normalizados del bloque de tamaño $p_n \times p_n$ de una matriz con distribución Haar sobre $\mathcal{G}_\beta(n)$ para $\beta = 1, 2$, converge débilmente a la ley del círculo.

Teorema 4.4. (*Ley del Círculo*) Sea Γ_n una matriz con distribución de Haar en $\mathcal{G}_\beta(n)$ con $\beta = 1, 2$ y \mathbf{U}_n su bloque superior izquierdo de tamaño $p_n \times p_n$. Sean $\lambda_1, \lambda_2, \dots, \lambda_p$ los eigenvalores de \mathbf{U}_n y μ_n la distribución empírica de $\sqrt{n/p}\lambda_i$. Denotemos por μ a la ley uniforme sobre $\{z \in \mathbb{C} : |z| \leq 1\}$. Si $p \rightarrow \infty$ y $p = o(\sqrt{n})$, entonces

$$\eta(\mu_n, \mu) \rightarrow 0$$

en probabilidad, cuando $n \rightarrow \infty$.

Demostración. Sin pérdida generalidad supongamos que $\Gamma_n \in \mathcal{U}(n)$. Sea \mathbf{Z}_n una matriz de tamaño $p_n \times p_n$ con entradas i.i.d. y ley gaussiana compleja estándar, por la Ley del Círculo (Véase Capítulo 1) si ν_n denota la distribución empírica de los eigenvalores de $\frac{1}{\sqrt{p_n}}\mathbf{Z}_n$, se cumple que

$$\eta(\nu_n, \mu) \rightarrow 0$$

en probabilidad, cuando $n \rightarrow \infty$, es decir para todo $\epsilon > 0$,

$$\lim_{n \rightarrow \infty} P(\eta(\nu_n, \mu) \geq \epsilon) = 0,$$

del Teorema 3.2 se sigue que

$$\lim_{n \rightarrow \infty} P(\eta(\mu_n, \mu) \geq \epsilon) - P(\eta(\nu_n, \mu) \geq \epsilon) = 0,$$

para todo $\epsilon > 0$, de lo cual se sigue el resultado deseado. ■

4.4. Ley de Marchenko-Pastur

Como se mencionó en la Sección 3 del primer capítulo, Marchenko y Pastur [49] estudiaron la convergencia de la función de distribución empírica de los eigenvalores de matrices de covarianza probando el Teorema 1.4.

El objetivo de esta sección es estudiar un resultado análogo considerando matrices con distribución de Haar truncadas. Esto es, dada U_n como en el Teorema 3.2 es de interés estudiar el comportamiento de la función de distribución empírica de los eigenvalores de la matriz $U_n^* U_n$, reescalados por constantes apropiadas, a la que denotaremos por F_{p_n, q_n} .

Antes de proseguir hay que recordar la definición de distancia de Lévy, así como introducir un lema que será de utilidad en la demostración del resultado principal de esta sección.

Definición 12. Si F_1 y F_2 son funciones de distribución, la distancia de Lévy entre ellas se define como

$$L(F_1, F_2) := \inf\{\epsilon > 0 : F_1(x - \epsilon) - \epsilon \leq F_2(x) \leq F_1(x + \epsilon) + \epsilon \text{ para todo } x \in \mathbb{R}\}.$$

Es bien conocido que esta distancia caracteriza la convergencia débil de las medidas de probabilidad correspondientes, es decir para cualquier sucesión de funciones de distribución $\{F, F_n, n \geq 1\}$ asegurar que

$$\lim_{n \rightarrow \infty} \int_{\mathbb{R}} g(x) dF_n(x) = \int_{\mathbb{R}} g(x) dF(x)$$

para toda función g continua y acotada, es equivalente a tener que

$$\lim_{n \rightarrow \infty} L(F_n, F) = 0,$$

que a su vez equivale a la convergencia puntual $F_n(x) \rightarrow F(x)$ en todos los puntos de continuidad de F . Por otra parte, el siguiente enunciado nos da una implicación más de tener convergencia en distancia de Levy.

Lema 4.5. Sea $\{F_n\}_{n \leq 1}$ una sucesión de funciones de distribución aleatorias y F una función de distribución continua dada. Si $L(F_n, F) \rightarrow 0$ en probabilidad, cuando $n \rightarrow \infty$, entonces

$$\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \rightarrow 0$$

en probabilidad.

Demostración. Como $L(F_n, F) \rightarrow 0$ en probabilidad, para cualquier ε existe $N \geq 1$ tal que

$$P(F(x - \varepsilon) - \varepsilon \leq F_n(x) \leq F(x + \varepsilon) + \varepsilon \quad \text{para todo } x \in \mathbb{R}) > 1 - \varepsilon$$

siempre que $n \geq N$.

De modo que si se cumple que

$$F(x - \varepsilon) - \varepsilon \leq F_n(x) \leq F(x + \varepsilon) + \varepsilon \quad \text{para todo } x \in \mathbb{R},$$

entonces

$$\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| \leq \sup_{|x-y| \leq \varepsilon} |F(x) - F(y)| + \varepsilon.$$

Por otra parte, como $F(+\infty) = 1$, $F(-\infty) = 0$ y $F(x)$ es continua, se tiene que es uniformemente continua y por tanto existe $\delta_0 > 0$ tal que

$$\sup_{|x-y| \leq \delta} |F(x) - F(y)| \leq \varepsilon$$

siempre que $\delta \in (0, \delta_0)$.

Así, para todo $\varepsilon \in (0, \delta_0)$ se tiene

$$P\left(\sup_{x \in \mathbb{R}} |F_n(x) - F(x)| < 2\varepsilon\right) > 1 - \varepsilon$$

siempre que $n \leq N$, lo cual nos da la convergencia en probabilidad que se quería. ■

Teorema 4.6. (Ley de Marchenko-Pastur) Sean Γ_n y \mathbf{Z}_n matrices aleatorias definidas como en el Teorema ??, para $\beta = 1, 2$. Considere \mathbf{U}_n el bloque superior izquierdo de tamaño $p_n \times q_n$ de Γ_n y F_{p_n, q_n} la distribución empírica de $(^{n/p_n})\lambda_i$, donde $\lambda_i, 1 \leq i \leq q$ son los eigenvalores de $\mathbf{U}_n^* \mathbf{U}_n$, es decir,

$$F_{p_n, q_n}(x) = \frac{1}{q_n} \sum_{i=1}^{q_n} \mathbf{1}_{((^{n/p_n})\lambda_i, \infty)}(x), \quad x \in \mathbb{R}. \quad (4.4)$$

Si $p_n \rightarrow \infty$, $\frac{q_n}{p_n} \rightarrow c > 0$ y $p_n = o(n/\log n)$, entonces

$$\sup_{x \in \mathbb{R}} |F_{p_n, q_n}(x) - F(x)| \rightarrow 0, \quad (4.5)$$

en probabilidad cuando $n \rightarrow \infty$, donde $F(x)$ es la ley de Marchenko Pastur, cuya densidad esta dada por (1.6).

Demostración. Note que por el lema anterior basta probar que $L(F_{p_n, q_n}, F) \rightarrow 0$ en probabilidad, cuando $n \rightarrow \infty$.

Por otra parte, sea \mathbf{Y}_n el bloque superior izquierdo de tamaño $p_n \times q_n$ de \mathbf{Z}_n y \tilde{F}_{p_n, q_n} la función de distribución empírica de los eigenvalores de $1/p_n \mathbf{Y}_n^* \mathbf{Y}_n$. Sean

$$\mathbf{A} = \frac{\sqrt{n}}{\sqrt{p_n}} \mathbf{U}_n^* \quad \text{y} \quad \mathbf{B} = \frac{1}{\sqrt{p_n}} \mathbf{Y}_n^*,$$

por el Lema C.6 (Véase Apéndice C) se obtiene la siguiente desigualdad

$$L^4(F_{p_n, q_n}, \tilde{F}_{p_n, q_n}) \leq \frac{2}{p_n^4} \left\{ \text{tr} [(\sqrt{n} \mathbf{U}_n - \mathbf{Y}_n)^* (\sqrt{n} \mathbf{U}_n - \mathbf{Y}_n)] \right\} \left\{ \text{tr}(n \mathbf{U}_n^* \mathbf{U}_n) + \text{tr}(\mathbf{Y}_n^* \mathbf{Y}_n) \right\}.$$

Note que considerando $\epsilon_n(m)$ como se definió en el Teorema ?? se sigue que

$$\text{tr} [(\sqrt{n} \mathbf{U}_n - \mathbf{Y}_n)^* (\sqrt{n} \mathbf{U}_n - \mathbf{Y}_n)] \leq p_n q_n \epsilon_n(q_n).$$

Por otra parte, recordando que $\|\mathbf{A}\|_M := \sqrt{\text{tr}(\mathbf{A}^* \mathbf{A})}$ es norma se tiene que

$$\begin{aligned} \text{tr}(n \mathbf{U}_n^* \mathbf{U}_n) &= \|\sqrt{n} \mathbf{U}_n\|_M^2 \\ &\leq (\|\sqrt{n} \mathbf{U}_n - \mathbf{Y}_n\|_M + \|\mathbf{Y}_n\|_M)^2 \\ &\leq 2p_n q_n \epsilon_n(q_n) + 2 \text{tr}(\mathbf{Y}_n^* \mathbf{Y}_n). \end{aligned}$$

De las desigualdades anteriores se obtiene

$$\begin{aligned} L^4(F_{p_n, q_n}, \tilde{F}_{p_n, q_n}) &\leq \frac{2}{p_n^4} p q \epsilon_n(q_n) \{2p_n q_n \epsilon_n(q_n) + 3 \text{tr}(\mathbf{Y}_n^* \mathbf{Y}_n)\} \\ &\leq \frac{2p_n \epsilon_n(q_n)}{p_n} \left(\frac{2q_n \epsilon_n(q_n)}{p_n} + \frac{3}{p_n^2} \text{tr}(\mathbf{Y}_n^* \mathbf{Y}_n) \right). \end{aligned}$$

Observemos que cuando $n \rightarrow \infty$, el segundo término dentro del paréntesis converge a $3c$ en probabilidad, por la Ley de grandes números. Por lo que solo resta mostrar $\epsilon_n(q_n)$ converge en probabilidad a cero. En efecto, por hipótesis se tiene que $p_n \rightarrow \infty$, $\frac{q_n}{p_n} \rightarrow c > 0$ y $p_n = o(n/\log n)$, por lo que es inmediato que $q_n = o(n/\log n)$ y del Teorema ?? se sigue que $\epsilon_n(q_n) \rightarrow 0$ en probabilidad cuando $n \rightarrow \infty$.

Con lo anterior se ha probado que $L(F_{p_n, q_n}, \tilde{F}_{p_n, q_n}) \rightarrow 0$ en probabilidad cuando $n \rightarrow \infty$, por otra parte del Teorema 1.4 se tiene la convergencia $L(\tilde{F}_{p_n, q_n}, F) \rightarrow 0$ en probabilidad. Ambas convergencias nos garantizan el resultado que se quería. ■

Apéndice A

Resultados sobre la Distribución Espectral de $Z'Z$

A lo largo de este apéndice considere $\mathbf{Z} = (z_{ij})_{1 \leq i \leq p, 1 \leq j \leq q}$ una matriz aleatoria con entradas complejas i.i.d. con distribución gaussiana estándar y denotemos por

$$\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n, \quad (\text{A.1})$$

a los eigenvalores de la matriz $\mathbf{Z}'\mathbf{Z}$.

Más aún, en algunos de los resultados se considerará una sucesión de matrices aleatorias $\{\mathbf{Z}_n\}_{n \geq 1}$ y para cada \mathbf{Z}_n los eigenvalores de $\mathbf{Z}'_n \mathbf{Z}_n$ serán denotados tal como en [A.1](#).

La primera afirmación del Lema siguiente se debe a Hsu [\[33\]](#) y nos proporciona una expresión para la densidad conjunta de, mientras que la segunda es simplemente el Jacobiano de la transformación que lleva a una matriz definida positiva y Hermitiana a sus eigenvalores y puede consultarse por ejemplo en [\[36\]](#).

Lema A.1. *Sea $\mathbf{Z} = (z_{ij})_{1 \leq i \leq p, 1 \leq j \leq q}$ una matriz de entradas complejas con función de densidad conjunta $g(\mathbf{Z}\mathbf{Z}')$, entonces la función de densidad conjunta de $\mathbf{R} = \mathbf{Z}\mathbf{Z}'$ está dada por*

$$\pi^{pq} (\Gamma_p(q))^{-1} (\det(\mathbf{R}))^{q-p} g(\mathbf{R}).$$

Más aún si $g(\mathbf{U}\mathbf{\Lambda}\mathbf{U}') = g(\mathbf{\Lambda})$ para toda \mathbf{U} en $\mathcal{O}(n)$, entonces la función de densidad conjunta de

$$\mathbf{\Gamma} = \text{diag}(\lambda_1, \dots, \lambda_p),$$

está dada por

$$\frac{\pi^{p(q+p-1)}}{\Gamma_p(q)\Gamma_p(p)}g(\Lambda)\left(\prod_{i=1}^p\lambda_i\right)^{p-q}\prod_{1\leq i<j\leq p}(\lambda_i-\lambda_j)^2,$$

donde $\Gamma_p(q) = \pi^{p(p-1)/2} \prod_{i=1}^p \Gamma(q-i+1)$.

Los siguiente resultados son bastante conocidos en teoría de matrices aleatorias y sus prueba pueden encontrarse en [2], la idea principal es utilizar el método de momentos para probar la convergencia débil. La técnica que consiste en valerse del método de momentos para demostrar la convergencia débil de distribuciones espectrales de matrices aleatorias o bien de funciones que dependen de los eigenvalores, es muy popular en teoría de matrices aleatorias. Esto se debe a que la distribución conjunta de la eigenvalores no tiene forma explícita, en general, mientras que un momento bajo la medida espectral es simplemente la traza de un potencia de la matriz, y por lo tanto es función explícita de la entradas.

Lema A.2. Sean $\{p_n\}_{n\geq 1}$ y $\{q_n\}_{n\geq 1}$ sucesiones de enteros positivos tales que $p_n \rightarrow \infty$ y $\frac{p_n}{q_n} \rightarrow \eta \in (0, \infty)$. Para cada entero n , considere \mathbf{Z}_n una matriz aleatoria de tamaño $p_n \times q_n$ con entradas i.i.d. con distribución gaussiana estándar, entonces para todo entero $k \geq 1$ se cumplen las siguientes afirmaciones:

i)

$$\mathbb{E}[\text{tr}(\mathbf{Z}'_n \mathbf{Z}_n)^k] \sim p_n^k q_n \sum_{r=0}^{k-1} \frac{1}{r+1} \left(\frac{q_n}{p_n}\right)^r \binom{k}{r} \binom{k-1}{r},$$

cuando $n \rightarrow \infty$.

ii)

$$\frac{\text{tr}((\mathbf{Z}'_n \mathbf{Z}_n)^k)}{q_n^{k+1}} \rightarrow \sum_{r=0}^{k-1} \frac{\eta^{k-r}}{r+1} \binom{k}{r} \binom{k-1}{r},$$

en probabilidad, cuando $n \rightarrow \infty$.

Lema A.3. Dado $\varepsilon \in (0, 1)$, sean $\{p_n\}_{n\leq 1}$ y $\{q_n\}_{n\leq 1}$ sucesiones de enteros positivos tales que $p_n \rightarrow \infty$ y $\varepsilon \leq \frac{p_n}{q_n} \leq \varepsilon^{-1}$ para toda $n \geq 1$. Para cada entero n considere \mathbf{Z}_n una matriz aleatoria de tamaño $p_n \times q_n$ con entradas i.i.d. con distribución gaussiana estándar y considere los eigenvalores como en A.1, entonces se cumplen las siguientes afirmaciones:

i) $\text{Var}[\text{tr}((\mathbf{Z}'_n \mathbf{Z}_n)^2)] \sim p_n^2 q_n^2 + 4p_n q_n (p_n + q_n)^2$, cuando $n \rightarrow \infty$.

ii) $\text{Cov}[\text{tr}(\mathbf{Z}'_n \mathbf{Z}_n), \text{tr}((\mathbf{Z}'_n \mathbf{Z}_n)^2)] \sim 2p_n q_n (p_n + q_n)$, cuando $n \rightarrow \infty$.

Apéndice B

Resultados sobre Cotas y Aproximaciones para Colas

Lema B.1. Sea ξ variable aleatoria con distribución normal estándar y considere

$$I(x) = \sup_{\theta \in \mathbb{R}} \{\theta x - \log(\mathbb{E}[\exp(\theta \xi^2)])\},$$

para $x \in \mathbb{R}$, entonces se cumplen las siguientes propiedades.

i) $\mathbb{E}[\exp(\theta \xi^2)] = (1 - 2\theta)^{-1/2}$ para todo $\theta < 1/2$.

ii)

$$I(x) = \begin{cases} \frac{(x-1-\log x)}{2}, & \text{si } x > 0; \\ +\infty, & \text{otro caso.} \end{cases}$$

iii) Sea $J(x) = I(x)/x$, para toda $x > 0$. Tanto $I(x)$ como $J(x)$ son crecientes en $(1, \infty)$ y decrecientes en $(0, 1]$.

Lema B.2. Sea Z_β una variable aleatoria con distribución $\mathbb{NF}_\beta(0, 1)$ con $\beta = 1, 2$ y 4 respectivamente. Se cumple que

$$\mathbb{P}(|Z_\beta| \geq x) \leq \varphi_\beta(x) e^{-\beta x^2/2} \text{ para toda } x > 0,$$

donde

$$\varphi_\beta(x) = \begin{cases} 1/x, & \text{si } \beta = 1; \\ 1, & \text{si } \beta = 2; \\ 2x^2 + 1, & \text{si } \beta = 4. \end{cases}$$

Demostración. En vista de la definición de $\mathbb{NF}_\beta(0, 1)$, se tiene que

$$\mathbb{P}(|\mathbb{NF}_\beta(0, 1)| \geq x) = \mathbb{P}\left(\frac{\sum_{i=1}^{\beta} \xi_i^2}{\beta} \geq x^2\right),$$

donde ξ_i son variables aleatorias i.i.d. con ley $N(0, 1)$.

Caso 1. Para $\beta = 1$, es bien conocida la desigualdad

$$P(|Z_1| \geq x) \leq \frac{2}{\sqrt{2\pi x}} e^{-x^2/2} \geq \frac{1}{x} e^{-x^2/2},$$

para toda $x > 0$.

Caso 2. Para $\beta = 2$ note que $(\xi_1^2 + \xi_2^2)/2 \stackrel{L}{=} \exp(1)$, luego

$$P(|Z_2| \geq x) \leq \frac{2}{\sqrt{2\pi x}} e^{-x^2/2} = \frac{1}{x} e^{-x^2},$$

para toda $x > 0$.

Caso 3. Para $\beta = 4$ se tiene que $\xi_1^2 + \xi_2^2 + \xi_3^2 + \xi_4^2 \stackrel{L}{=} \chi^2(4)$, luego

$$P(|Z_4| \geq x) = \int_{4x^2}^{\infty} \frac{t}{4} e^{-t/2} dt = \int_{2x^2}^{\infty} y e^{-y} dy = (2x^2 + 1) e^{-2x^2},$$

para toda $x > 0$.

Combinando los tres casos se obtiene la desigualdad deseada. ■

Lema B.3. Sea $\{\xi\}_{i \geq 1}$ una sucesión de variables aleatorias i.i.d. con distribución gaussiana estándar, para cualesquiera n, n con $n \geq m \geq 3$ sea

$$W = \left(\frac{\sum_{i=1}^m \xi_i^2}{\sum_{i=1}^n \xi_i^2} \right)^{1/2},$$

se cumplen las siguientes igualdades

i) $E[W^{-2}] \leq n$.

ii) Para todo $a > 0$,

$$E[e^{-a/w^2}] \leq 3 \left(1 + \frac{2a}{3(m + \sqrt{2an})} \right)^{-n/2}.$$

Demostración. i) El resultado es obvio para el caso $m = n$, por lo cual se supondrá $3 \leq m < n$. Reescribiendo W^{-2} como

$$W^{-2} = 1 + \frac{\sum_{i=m+1}^n \xi_i^2}{\sum_{i=m+1}^m \xi_i^2},$$

entonces

$$\begin{aligned}
 E[W^{-2}] &= 1 + E \left[\frac{\sum_{i=m+1}^n \xi_i^2}{\sum_{i=1}^m \xi_i^2} \right] \\
 &= 1 + E \left[\sum_{i=m+1}^n \xi_i^2 \right] E \left[\frac{1}{\sum_{i=1}^m \xi_i^2} \right] \\
 &\leq 1 + (n - m) E \left[\frac{1}{\sum_{i=1}^3 \xi_i^2} \right],
 \end{aligned}$$

al calcular la esperanza del último término se tiene,

$$\begin{aligned}
 E \left[\frac{1}{\sum_{i=1}^3 \xi_i^2} \right] &= \frac{1}{(\sqrt{2\pi})^3} \int_{\mathbb{R}^3} \frac{\exp\{-\frac{x^2+y^2+z^2}{2}\}}{x^2 + y^2 + z^2} dx dy dz \\
 &= \frac{1}{(\sqrt{2\pi})^3} \int_0^\infty \int_0^\pi \int_0^{2\pi} \frac{\exp\{-\frac{r^2}{2}\}}{r^2} r^2 \sin \theta_1 dr d\theta_1 d\theta_2 \\
 &= \frac{1}{(\sqrt{2\pi})^3} \frac{1}{2} \int_{-\infty}^\infty \exp\left\{-\frac{r^2}{2}\right\} dr \int_0^\pi 2\pi \sin \theta_1 d\theta_1 \\
 &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^\infty \exp\left\{-\frac{r^2}{2}\right\} dr \\
 &= 1,
 \end{aligned}$$

entonces

$$E[W^{-2}] \leq 1 + (n - m)$$

de donde claramente se sigue el resultado.

ii) Por la independencia de $\sum_{i=1}^m \xi_i^2$ con las variables $\xi_{m+1}, \xi_{m+2}, \dots, \xi_n$ se tiene que

$$E[e^{-aW^{-2}}] = e^{-a} E[M^{n-m}],$$

donde

$$M = E \left[\exp \left\{ -\frac{a\xi_n^2}{\sum_{i=1}^m \xi_i^2} \right\} \mid \xi_1, \dots, \xi_m \right].$$

Por otra parte, se sabe que $E[\exp(-\beta\xi_n^2)] = (1 + 2\beta)^{-1/2}$ para $\beta > -1/2$, así que

$$M = \left(1 + 2a \left(\sum_{i=1}^m \xi_i^2 \right)^{-1} \right)^{-1/2},$$

y por lo tanto se llega a que

$$E[e^{-aW^{-2}}] = e^{-a} E \left\{ \left(1 + \frac{2a}{\sum_{i=1}^m \xi_i^2} \right)^{-(n-m)/2} \right\}$$

■

Apéndice C

Resultados Técnicos

Lema C.1. Sea $\Gamma(x)$, $x > 0$ la función Gamma, definida como

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt.$$

Se cumplen las siguientes desigualdades.

i) $1 - \frac{1}{6n} < \frac{\Gamma(n + (1/2))}{\sqrt{n}\Gamma(n)} < 1$ para toda $n \geq 1$.

ii) $\left| \frac{\Gamma((n+1)/2)}{\sqrt{n/2}\Gamma(n/2)} - 1 \right| < \frac{3}{5n}$ para toda $n \geq 1$.

El siguiente resultado nos da la identidad conocida como integral de Selberg cuya prueba puede encontrarse en diversas fuentes, como en [25] por mencionar una.

Lema C.2 (Integral de Selberg). Sea $n \leq 2$ un entero, α, β y γ reales positivos, entonces

$$\begin{aligned} & \int_0^1 \int_0^1 \cdots \int_0^1 \prod_{i=1}^n x_i^{\alpha-1} (1-x_i)^{\beta-1} \prod_{1 \leq j < k \leq n} |x_j - x_k|^{2\gamma} dx_1 dx_2 \cdots dx_n \\ &= \prod_{l=0}^{n-1} \frac{\Gamma(1 + \gamma + l\gamma) \Gamma(\alpha + l\gamma) \Gamma(\beta + l\gamma)}{\Gamma(1 + \gamma) \Gamma(\alpha + \beta + (n+l-1)\gamma)}. \end{aligned}$$

Lema C.3. Sea $f(u, v)$ una función real, tal que sus tres derivadas de segundo orden existen y son acotadas en valor absoluto por una constante M , sobre $[a, b] \times [c, d]$. Entonces se cumple la siguiente igualdad,

$$\begin{aligned} \frac{1}{n^2} \sum_{j=j_1}^{j_2} \sum_{i=i_1}^{i_2} f\left(\frac{i}{n}, \frac{j}{n}\right) &= \int_{\frac{j_1}{n}}^{\frac{j_2+1}{n}} \int_{\frac{i_1}{n}}^{\frac{i_2+1}{n}} f(x, y) dx dy \\ &- \frac{1}{2n^3} \sum_{j=j_1}^{j_2} \sum_{i=i_1}^{i_2} f'_x\left(\frac{i}{n}, \frac{j}{n}\right) - \frac{1}{2n^3} \sum_{j=j_1}^{j_2} \sum_{i=i_1}^{i_2} f'_y\left(\frac{i}{n}, \frac{j}{n}\right) + \epsilon, \end{aligned} \tag{C.1}$$

donde

$$|\epsilon| \leq \frac{(i_2 - i_1)(j_2 - j_1)M}{n^4},$$

para cualesquiera i_1, i_2, j_1 y j_2 tales que $na \leq i_1 < i_2 \leq nb - 1$ y $nc \leq j_1 < j_2 \leq nd - 1$.

Lema C.4. Las siguientes desigualdades se satisfacen:

i) Para todo $x \in (0, 1]$ se cumple que

$$x - 1 - \log x \geq \frac{(x - 1)^2}{2}.$$

ii) Para todo $x \in (0, 1/4]$ se cumple que

$$2x - \log(1 + 2x) \geq x^2.$$

iii) Para todo $x \in (0, 1/4]$ se cumple que

$$(1 - x)^{-2} \geq 1 + 2x$$

y que

$$(1 + x)^{-2} \leq 1 - x.$$

Lema C.5. Si \mathbf{A} y \mathbf{B} son matrices de tamaño $n \times n$ con eigenvalores complejos $\lambda_1, \dots, \lambda_n$ y τ_1, \dots, τ_n , respectivamente, entonces se cumple que

$$\max_i |\lambda_i - \tau_i| \leq 2^{2-1/n} (\|\mathbf{A}\|_2 + \|\mathbf{B}\|_2)^{1-1/n} \|\mathbf{A} - \mathbf{B}\|_2^{1/n}$$

donde $\|\cdot\|_2$ denota la norma usual de matrices.

Lema C.6. Sean \mathbf{A} y \mathbf{B} matrices aleatorias de tamaño $p \times n$ con entradas complejas. Si denotamos por $F^{\mathbf{A}\mathbf{A}^*}$, $F^{\mathbf{B}\mathbf{B}^*}$ a las funciones de distribución empírica de $\mathbf{A}\mathbf{A}^*$ y $\mathbf{B}\mathbf{B}^*$ respectivamente, entonces

$$L^4(F^{\mathbf{A}\mathbf{A}^*}, F^{\mathbf{B}\mathbf{B}^*}) \leq \frac{2}{p^2} \text{tr}((\mathbf{A} - \mathbf{B})(\mathbf{A} - \mathbf{B})^*) \text{tr}(\mathbf{A}\mathbf{A}^* + \mathbf{B}\mathbf{B}^*),$$

donde L denota la distancia de Levy.

Bibliografía

- [1] Bai, Z. D. and Yin, Y.Q. (1993), Limit of the smallest eigenvalue of a large dimensional sample covariance matrix, *Annals of Probability* **21**, 1275 – 1294.
- [2] Bai, Z. D. (1997), Circular Law, *Annals of Probability* **25**, 194 – 529. [1.2.1](#), [A](#)
- [3] Bai, Z. D. (1999), Methodologies in spectral analysis of large dimensional random matrices, a review, *Statistica Sinica* **9**, 611 – 677. [4.1](#)
- [4] Bai, Z. D. and Silverstein, J. W. (2004), CLT for linear spectral statistics of large dimensional sample covariance matrix, *Annals of Probability* **32**, 553 – 605. [3.2.1](#)
- [5] Bjorck, W. (1994), Numerics of Gram-Schmidt orthogonalization, *Linear Algebra and its Applications* **197,198**, 279 – 316.
- [6] Bottcher, A. and Silbermann, B., (1999), *Introduction to Large Truncated Toeplitz Matrices*, Springer-Verlag: New York. ([document](#))
- [7] Borel, E. (1906), *Introduction Géométrique a Quelques Théories Physiques*, Gauthier-Villars, Paris. ([document](#)), [3.1](#)
- [8] Businger, P., Golub, G.H. in Wilkinson-Reinsch, *Linear Algebra*, Springer Verlag. [2.1.1](#)
- [9] Collins, B. (2003), *Intégrales Matricielles et Probabilités Non-commutatives*, Thèse de Doctorat of Université Paris 6. [3.1](#), [3.2.1](#)
- [10] Conrey, J. B., Farmer, D. W., Keating, J. P., Rubinstein, M.O. and Snaith, N.C. (2005), Integral moments of L-functions, *Proceedings of the London Mathematical Society* **91**, 33 – 104. [2](#)
- [11] D' Aristotle, A., Diaconis, P. and Newman, C. M. (2002). Brownian motion and the classical groups. *Probability, Statistics and Their Applications: Papers in Honor of Rabi Bhattacharya* 97 – 116. IMS Lecture Notes Monogr. Ser. Edited by K. Athreya, et al. 97 – 116. Beechwood, OH: Institute of Mathematical Statistics. ([document](#))

- [12] Dembo, Z. and Zeitouni, O. (1998), *Large Deviations Techniques and Applications*, Springer-Verlag, Second Edition.
- [13] Diaconis, P. and Freedman, D. (1987), A dozen de Finetti-style results in search of a theory, *Annales de l'Institut Henri Poincaré* **23**, 397 – 423. [3.1](#), [3.2](#)
- [14] Diaconis, P. and Shahshahni, M. (1987), The subgroup algorithm for generating uniform random variables, *Probability in the Engineering and Informational Sciences* **1**, 15 – 32. [2](#)
- [15] Diaconis, P., Eaton, M. and Lauritzen, L. (1992), Finite de Finetti theorems in linear models and multivariate analysis, *Scandinavian Journal of Statistics* **19**, 189 – 315. [3.1](#)
- [16] Diaconis P. and Shahshahani, M. (1994), On the eigenvalues of random matrices, *Journal of Applied Probability* **31**, 49 – 62. ([document](#))
- [17] Diaconis P. (2003), Patterns in eigenvalues: the 70th Josiah Willard Gibbs lecture, *Bulletin of the American Mathematical Society* **40**, 155 – 178. ([document](#)), [3.1](#)
- [18] Diaconis P., Coram M. (2003), New test of the correspondence between unitary eigenvalues and the zeros of Riemann's zeta function, *Journal of Physics A: Mathematical and General* **36**, 2883-2906. ([document](#))
- [19] Dueñez, E., Farmer, D. W., Froehlich, S., Hughes, C. P., Mezzadri, F. and Phan, (2010) Roots of the derivative of the Riemann zeta function and of characteristic polynomials, *Nonlinearity* **23**, 2599-2621.
- [20] Dyson, F. J. (1962), Statistical theory of the energy levels of complex systems I, *Journal of Mathematical and Physical Sciences* **3**, 140 – 156. ([document](#))
- [21] Dyson, F. J. (1962), Statistical theory of the energy levels of complex systems II, *Journal of Mathematical and Physical Sciences* **3**, 157 – 165. ([document](#))
- [22] Eaton M. L. (1983), *Multivariate Statistics: A Vector Approach*, Wiley, New York. [2.1](#)
- [23] Edelman A. and Rao N. R. (2005), Random matrix theory, *Acta Numerica* **14**, 233 – 297. [2.1.3](#)
- [24] Edelman, A. (1997), The probability that a random real gaussian matrix has k real eigenvalues, related distributions, and the circular law. *Journal of Multivariate Analysis* **60**, 203 – 232. [1.2.1](#)
- [25] Forrester, P. (2007), *Log-gases and Random Matrices*, London Mathematical Society Monographs.

- [26] Forrester, P. and Rains, E. (2008), Inter-Relationships Between Orthogonal, Unitary and Symplectic Matrix Ensembles, *MSRI Publications* **40**, 171. ([document](#))
- [27] Geman, S. (1986), The spectral radius of large random matrices. *Annals of Probability* **14**, 1318 – 1328. [2.2](#), [4.1](#)
- [28] Ginibre, J. (1965), Statistical ensembles of complex, quaternion and real matrices, *Journal of Mathematical Physics* **6**, 440 – 449. [1.1.2](#), [1.2.1](#)
- [29] Girko, V. L. (1984), Circle law, *Theory of Probability and its Applications* **4**, 694 – 706. [1.2.1](#)
- [30] Girko, V. L. (1984), On the circle law, *Theory of Probability and Mathematical Statistics* **28**, 15 – 23. [1.2.1](#)
- [31] Haar, A. (1933), Der Massbegriff in der Theorie der kontinuierlichen Gruppen, *Annals of Mathematics* **34**, 147 – 169. [1.4](#)
- [32] Heiberger, R. M. (1978), Algorithm AS127. Generation of random orthogonal matrices, *Applied Statistics* **27**, 199 – 206. [2](#)
- [33] Hsu, P. L. (1939), On the distribution of the roots of certain determinantal equations, *Annals of Eugenics* **9** 250 – 258. [A](#)
- [34] Halmos, P. (1974), *Measure Theory*, Springer-Verlag, New York. [1.4](#)
- [35] Hwang, C. R. (1986), A brief survey on the spectral radius and the spectral distribution of large dimensional random matrices with iid entries. In *Random Matrices and Their Applications: Contemporary Mathematics* 50, J. E. Cohen, H. Kesten, and C. M. Newman, editors, 145 – 152. AMS, Providence, RI. [1.2.1](#)
- [36] James, A. T. (1964), Distributions of matrix variates and latent roots derived from normal samples, *The Annals of Mathematical Statistics* **35**, 475 – 501. [A](#)
- [37] Jiang, T. (2005), Maxima of entries of Haar distributed matrices, *Probability Theory and Related Fields* **131**, 121 – 144. ([document](#)), [3.1](#), [3.2](#)
- [38] Jiang, T. (2006), How many entries of a typical orthogonal matrix can be approximated by independent normals?, *Annals of Probability* **34**, 1497 – 1529. ([document](#)), [3.3](#)
- [39] Jiang, T. (2009), Approximation of Haar distributed matrices and limiting distributions of eigenvalues of Jacobi ensembles, *Probability Theory and Related Fields* **144**(1), 221 – 246. [3.2](#), [4.3](#)

- [40] Jiang, T. (2010), The Entries of Haar-invariant matrices from the classical compact groups, *Journal of Theoretical Probability* **23**, 1227 – 1243. ([document](#))
- [41] Johansson, K. (1997), On random matrices from the compact classical groups, *The Annals of Mathematics* **145**, 519 – 545. ([document](#)), [4.2](#)
- [42] Johnstone, I. M. (2001), On the distribution of the largest eigenvalue in principal components analysis, *Annals of Statistics* **29**, 295 – 327. [1.2.2](#), [4.2](#)
- [43] Jonsson, D. (1982), Some limit theorems for the eigenvalues of a sample covariance matrix, *Journal of Multivariate Analysis* **12**, 1 – 38.
- [44] Katz, N. M. and Sarnak, P. (1999), Random Matrices, Frobenius Eigenvalues, and Monodromy, *AMS Colloquium Publications* 45, AMS, Providence, RI. [2](#)
- [45] Keating, J. and Snaith, N. (2000), Random Matrix Theory and $\zeta(\frac{1}{2} + it)$, *Communications in Mathematical Physics* **214**, 57 – 89. ([document](#)), [2](#)
- [46] Keating, J. and Snaith, N. (2000), Random matrix theory and L-functions at $s = \frac{1}{2}$, *Communications in Mathematical Physics* **214**, 91 – 110. ([document](#)), [2](#)
- [47] Khokhloy, V. (2006), The uniform distribution on a sphere in R^s , *Society for Industrial and Applied Mathematics* **50**, 389 – 399.
- [48] Li, S. and Teng, C. (2001), Non-central quaternion χ^2 ; t and F distributions and their properties, *Journal of Mathematical Study* **34**, 170 – 173. [3.3.2](#)
- [49] Marchenko, V.A. and Pastur, L. A. (1967), Distribution of some sets of random matrices, *Mathematics of the USSR, Sbornik. American Mathematical Society* **1**, 457 – 483. ([document](#)), [1.3.1](#), [4.4](#)
- [50] Mehta, M. L. (1960), On the statistical properties of the level spacings in nuclear spectra, *Nuclear Physics* **18**, 395 – 419. ([document](#)), [1.2.1](#)
- [51] Mezzadri, F. (2007), How to generate random matrices from the classical compact groups?, *Notices of the AMS* **54**, 592 – 604. [2.1.3](#)
- [52] Montgomery, H.L. (1973), The pair correlation of zeros of the Riemann zeta-function, *Proceedings of Symposia in Pure Mathematics* **24**, 181 – 93. [2](#)
- [53] Odlyzko, A. (1989), The 10^{20} th zero of the Riemann zeta-function and 70 million of its neighbors, Unpublished. [2](#)

- [54] Pan, G. M. and Zhou, W. (2007), Circular law, extreme singular values and potential theory, *Journal of Multivariate Analysis* **101**, 645 – 656. [1.2.1](#)
- [55] Petz, D., Réffy, J. (2004), On asymptotics of large Haar distributed unitary matrices, *Periodica Mathematica Hungarica* **49**, 103 – 117.
- [56] Petz, D., Réffy, J. (2005), Large deviations for the empirical eigenvalue density of truncated Haar unitary matrices, *Large deviations for the empirical eigenvalue density of truncated Haar unitary matrices* **133**, 175 – 189. [4.3](#)
- [57] Rains, E. M. (1997), High powers of random elements of compact Lie groups, *Probability Theory and Related Fields* **107**, 219 – 241. ([document](#))
- [58] Rubinstein, M. O. (1998), *Evidence for a spectral interpretation of the zeros of L-functions*, Thesis, Princeton University. [2](#)
- [59] Rutishauser, H. (1967), *Description of Algol 60*, Springer. [2.1.1](#)
- [60] Silverstein, J. W. (1979), On the randomness of eigenvectors generated from networks with random topologies, *SIAM Journal of Applied Mathematics* **37**, 235 – 245. [1.4](#)
- [61] Silverstein, J. W. (1984), Comments on a result of Yin, Bai and Krishnaiah, *Journal of Multivariate Analysis* **15**, 408 – 409. [4.1](#)
- [62] Sloane, N. (1983), Encrypting by Random Rotations, *Technical Memorandum, Bell Laboratories*. ([document](#))
- [63] Stam, A. J. (1982), Limit theorems for uniform distributions on spheres in highdimensional Euclidean spaces, *Journal of Applied Probability* **19**, 221 – 228. [3.1](#)
- [64] Stewart G.W. (1980), The efficient generation of random orthogonal matrices with an application to condition estimators, *Society for Industrial and Applied Mathematics: Journal of Numerical Analysis* **17**, 403 – 409. [2](#)
- [65] Tanner Martin A. and Thisted Ronald A. (1982), Remark AS R42: A remark on AS 127. Generation of random orthogonal matrices, *Journal of the Royal Statistical Society, (Applied Statistics)* **31**, 190 – 192. [2](#)
- [66] Tao, T. and Vu, V.(2008), Random matrices: the circular law, *Communications in Contemporary Mathematics* **10**, 261 – 307. [1.2.1](#)
- [67] Tracy, C. A. and Widom, H. (1994), Level-spacing distributions and Airy kernel, *Communications in Mathematical Physics* **159**, 151 – 174. ([document](#)), [1.2.2](#)

- [68] Tracy, C. A. and Widom, H. (1996), On the orthogonal and symplectic matrix ensembles, *Communications in Mathematical Physics* **177**, 727 – 754. [1.2.2](#)
- [69] Tracy, C.A. and Widom, H. (2000), The distribution of the largest eigenvalue in the Gaussian ensembles, *Calogero-Moser-Sutherland Models. Vol. 4 of CRM Series in Mathematical Physics*, Springer, Berlin, 461 – 472. [1.2.2](#)
- [70] Wedderburn R. W. M. (1975), Generating random rotations. Research report, *Rothamsted Experimental Station*. [2](#)
- [71] Wishart, J. (1928), The generalized product moment distribution in samples from a normal multivariate population, *Biometrika* **20**, 32 – 52. ([document](#))
- [72] Wigner, E. P. (1951), On the statistical distribution of the widths and spacing of nuclear resonance levels, *Mathematical Proceedings of the Cambridge Philosophical Society* **47**, 790 – 798. ([document](#))
- [73] Wigner, E. P. (1955), Characteristic vectors bordered matrices with infinite dimensions, *Annals of Mathematics* **62**, 548 – 564. ([document](#))
- [74] Wigner, E. P. (1957), Statistical properties of real symmetric matrices with many dimensions, *Canadian Mathematical Congress Proceeding*, Toronto 1957, pag. 174 – 184. ([document](#))
- [75] Wigner, E. P. (1958), On the distribution of the roots of certain symmetric matrices, *Annals of Mathematics* **67**, 325 – 327. ([document](#))
- [76] Wigner, E. P. (1962), Distribution laws for the roots of a random Hermitian matrix, *Statistical Theories of Spectra: Fluctuations* edited by C.E. Porter, Academic, New York. ([document](#))
- [77] Yin, Y. Q., Bai, Z. D. and Krishnaiah, P.R. (1983), Limiting behavior of the eigenvalues of multivariate F matrix, *Journal of Multivariate Analysis* **13**, 508 – 516. [4.1](#)
- [78] Yin, Y. Q., Bai, Z. D. and Krishnaiah, P.R. (1988), On the limit of the largest of the eigenvalues of the large dimensional sample covariance matrix, *Probabability Theory and Related Fields* **78**, 509 – 521. [4.1](#)
- [79] Yin, Y. Q. (1986), LSD for a class of random matrices, *Journal of Multivariate Analysis* **20**, 50 – 68.
- [80] Zyczkowski, K. and Kus, M. (1994), Random unitary matrices, *Journal of Physics A: Mathematical and General* **27**, 4235 – 4245.
- [81] Zyczkowski, K. and Sommers, H. J. (2000), Truncations of random unitary matrices, *Journal of Physics* **33**, 2045 – 2057. [4.3](#)