



CIMAT

Centro de Investigación en Matemáticas, A.C.

SOBRE EL ESPECTRO DE MATRICES LAPLACIANAS

T E S I S

Que para obtener el grado de
Maestro en Ciencias
con Orientación en
Probabilidad y Estadística

Presenta

Santiago Arenas Velilla

Director de Tesis:

Dr. Víctor Pérez-Abreu Carrión

Autorización de la versión final

Guanajuato, Gto., 6 de Julio de 2018

*Dedicado a mis padres y mi hermana,
sin su apoyo esto no sería posible.*

Agradecimientos

La realización de esta tesis fue todo un reto a nivel personal y profesional. No me queda más que agradecer enormemente a mi asesor el Dr. Víctor Pérez-Abreu Carrión por aceptar dirigir la tesis, por todas sus enseñanzas en los dos cursos que recibí y por todo el apoyo y confianza que me ha brindado a lo largo de estos semestres. Esto no sería posible sin toda su ayuda y dirección.

Doy gracias a mis sinodales el Dr. Octavio Arizmendi Echegaray y al Dr. Carlos Vargas Obieta por sus correcciones y sugerencias para lograr que la tesis saliera lo mejor posible. También al Dr. Juan Carlos Pardo Millán por su apoyo durante toda la maestría, desde la entrevista de admisión, así como su tutoría y el par de cursos que impartió, por todo esto, muchas gracias.

Agradezco a mis compañeros de la oficina D620, Adrian, Alberto, Camilo y Pablo (al menos durante el primer año), gracias a ustedes mi estancia en Cimat fue mucho más cómoda y agradable. También agradezco a mis compañeros de Maestría, Fernanda, Sonny, Sebastian, Mafer, Daniel³, Javier al igual que amigos de otros programas como Carlitos y Delio. Muchas gracias por su amistad y apoyo.

También agradezco a mis padres, Pablo y Angela y a mi hermana Paula, que

desde la distancia siempre me estuvieron apoyando y dándome ánimos y buenas energías para continuar, al igual que a mis amigos Jonathan, Maria Fernanda, Estefanía y Walter, su apoyo siempre fue muy importante.

Finalmente, quiero agradecer a CONACYT por la beca de maestría que me fue otorgada al igual que al Centro de Investigación en Matemáticas A.C (CIMAT A.C) por todas las facilidades y apoyos para participar en eventos.

Índice general

| | |
|--|-----------|
| Agradecimientos | II |
| 1. Preliminares | 1 |
| 1.1. Elementos de gráficas y matrices asociadas | 1 |
| 1.2. Propiedades de las matrices laplacianas | 5 |
| 1.3. Programación semidefinida | 9 |
| 1.4. Programación semidefinida en $\{-1, 1\}^n$ | 11 |
| 2. Sobre el espectro asintótico del laplaciano | 14 |
| 2.1. Distribución empírica espectral | 15 |
| 2.2. Densidad de la distribución asintótica espectral | 17 |
| 2.3. Mayor eigenvalor | 22 |
| 2.4. Estimación del mayor eigenvalor | 27 |
| 2.5. Sobre la distribución asintótica del mayor eigenvalor | 34 |
| 3. Aplicaciones | 40 |
| 3.1. \mathbb{Z}_2 sincronización con ruido general | 40 |
| 3.2. Sincronización sobre \mathbb{Z}_2 | 44 |
| 3.2.1. Solución del problema | 47 |
| 3.2.2. Demostración del teorema principal | 48 |
| 3.2.3. Estimación del mayor eigenvalor | 50 |

| | |
|---|-----------|
| <i>ÍNDICE GENERAL</i> | V |
| 3.3. Modelo de bloques estocásticos | 51 |
| 3.3.1. Solución del problema | 54 |
| 3.3.2. Demostración del teorema principal | 55 |
| A. Reporte de simulaciones | 66 |
| A.1. Caso Normal | 67 |
| A.2. Caso uniforme | 68 |
| A.3. Caso Bernoulli | 69 |
| Bibliografía | 69 |

Introducción

Desde los trabajos pioneros de Wigner en los 50's, el estudio del espectro asintótico de diversos tipos de matrices aleatorias ha sido un tema de interés con aplicaciones en diversos campos como la física, la biología, las telecomunicaciones y las finanzas. Asimismo, dentro de las matemáticas se han considerado las matrices aleatorias asociadas a gráficas aleatorias como las de adyacencia y laplaciana.

Entre muchos otros aspectos, se ha considerado el estudio de la distribución asintótica de la distribución empírica espectral, así como el comportamiento asintótico del mayor eigenvalor para una diversidad de matrices aleatorias, obteniéndose resultados universales.

Esta tesis considera matrices laplacianas generales, en el sentido de que no provienen necesariamente de una gráfica. Para estas matrices la distribución asintótica empírica espectral fue considerada en Bryc-Dembo-Jiang [8] obteniéndose ésta como la convolución libre de una distribución semicircular con una distribución normal. La universalidad de este resultado fue considerada en Ding-Jiang [9], quienes también abordan el estudio asintótico del mayor eigenvalor de una matriz laplaciana aleatoria.

En un trabajo reciente (del 2018), Bandeira [6] estudia problemas de optimización convexa en los que usa la técnica de relajación para obtener problemas de programación semidefinida en donde el comportamiento asintótico del ma-

yor eigenvalor de una matriz laplaciana determina la unicidad de la solución del problema de optimización. Asimismo, el trabajo [6] considera la estimación del mayor eigenvalor mediante el máximo elemento de la diagonal del laplaciano, lo cual es muy conveniente desde el punto de vista computacional.

El objetivo principal de esta tesis es mostrar cómo los resultados en [9] pueden ser usados para dar un marco general a varios de los resultados y aplicaciones en [6]. Este enfoque permite dar demostraciones diferentes, así como obtener extensiones de algunos resultados y plantear varias conjeturas.

Más específicamente, estudiamos el comportamiento del mayor eigenvalor de matrices laplacianas aleatorias generales, con los que obtuvimos varias consecuencias interesantes e importantes que presentamos en el Capítulo 2.

Una ventaja del enfoque que presentamos es que nos permite desarrollar pruebas más sencillas desde una metodología diferente a la presentada en [6], aunque una posible dificultad es que en algunos casos es necesario encontrar cotas ajustadas. Lo anterior nos lleva a estudiar desviaciones grandes y a calcular de forma exacta ciertas desigualdades.

También presentamos un estudio de simulación, con el cual obtuvimos intuición para plantear una conjetura incluida en el Capítulo 2.

La organización y presentación de la tesis y nuestros resultados son los siguientes.

El Capítulo 1 contiene los preliminares de teoría de gráficas, matrices y optimización convexa necesarios para el desarrollo de la tesis. En la Sección 1.1 se mencionan los conceptos de matriz de adyacencia y laplaciana asociadas a gráficas, así como la noción de matriz laplaciana general y los principales resultados acerca de sus eigenvalores cuando la dimensión de la matriz está fija. En particular se muestra el rol del mayor eigenvalor de un laplaciano y su relación con el segundo eigenvalor de otro laplaciano (Proposición 1.2), resultado esencial para

los problemas de optimización del Capítulo 3. En la Sección 1.2 se resumen los principales elementos de optimización convexa con énfasis en la programación semidefinida. En la Sección 1.3 se presenta un ejemplo de programación semidefinida en $\{-1, 1\}^n$, cuya solución óptima existe cuando el segundo valor de cierto laplaciano es positivo. Este ejemplo es el marco general para estudiar la aplicaciones del Capítulo 3.

El Capítulo 2 trata sobre resultados asintóticos sobre la distribución empírica espectral y el mayor eigenvalor de una matriz laplaciana aleatoria. En la Sección 2.1 se presentan resultados de Ding-Jiang [9] sobre la distribución asintótica empírica espectral de matrices de adyacencia y laplacianas con entradas con distinta distribución, trabajo que se inspira en el artículo pionero de Bryc-Dembo-Jiang [8]. En el caso del laplaciano esta distribución límite es la convolución libre de una distribución semicircular con una distribución normal. En la Sección 2.2 presentamos los resultados de un estudio de simulación que nos llevan a conjeturar que esta distribución límite pudiera ser aproximada por una distribución mezcla de la distribución del semicírculo y la distribución normal. En la Sección 2.3 se enuncian resultados de [9] sobre la convergencia del mayor eigenvalor de matrices laplacianas (Teorema 2.4). Como una de sus consecuencias proponemos un enfoque alternativo al estudio de estimar el mayor eigenvalor de una matriz laplaciana que hace Bandeira en [6], lo cual nos permite obtener extensiones que dan un planteamiento unificado a las aplicaciones presentadas en el Capítulo 3. Asimismo, como otra consecuencia del Teorema 2.4, se obtuvo la convergencia del mayor eigenvalor en modelos de bloques estocásticos. En la Sección 2.4 se presenta la estimación del mayor eigenvalor mediante el elemento máximo de la diagonal de la matriz laplaciana, lo cual es útil para reducir el costo computacional del cálculo de los eigenvalores. Para esto obtuvimos otra consecuencia del Teorema 2.4 que nos permite proponer un enfoque alterno y unificado a aplica-

ciones consideradas en [6] y dar pruebas alternativas a sus resultados; como se muestra en el Capítulo 3. Finalmente, en la Sección 2.5 se presenta un resultado que obtuvimos sobre la distribución asintótica del mayor eigenvalor de matrices laplacianas que provienen de matrices simétricas con entradas normales estándar independientes, lo cual es distribución Gumbel.

El Capítulo 3 trata sobre tres aplicaciones a problemas de programación semidefinida donde el mayor eigenvalor de matrices laplacianas surge de manera natural para la solución de los mismos. Estas aplicaciones aparecen en un artículo de Bandeira [6]. Nuestra metodología es diferente a la presentada en [6], la que nos permite dar pruebas más sencillas y en algunos casos extensiones de sus resultados. Nuestro enfoque se basa principalmente en la aplicación de las consecuencias del Teorema 2.4 sobre el comportamiento del mayor eigenvalor asintótico de matrices laplacianas que discutimos en el Capítulo 2 junto con los resultados de optimización convexa del Capítulo 1. En la Sección 3.1 se presenta el problema denominado *Sincronización con ruido general*, en el que obtuvimos, con una demostración más sencilla, condiciones para garantizar que este problema tenga solución única con alta probabilidad (Proposición 3.1) y además obtuvimos cotas, superior e inferior, para el mayor eigenvalor de una matriz laplaciana con entradas normal estándar (Proposición 3.2). En la Sección 3.2 se considera el problema de *Sincronización sobre \mathbb{Z}_2* , en el que obtuvimos una extensión a los resultados presentados en [6] usando nuestro enfoque. Finalmente, en la Sección 3.3 se estudia el Modelo de bloques estocásticos con 2 comunidades, en el cual usamos nuestro resultado para matrices a bloques (Proposición 2.5), resultados sobre el comportamiento del mayor eigenvalor de matrices de adyacencia (Lema 2.1) y técnicas de desviaciones grandes (Proposición 3.3), para obtener estimaciones adecuadas para el mayor eigenvalor de ciertas matrices laplacianas asociadas al modelo de bloques estocásticos con dos comunidades. Con todos los elementos

anteriores, dimos una prueba alternativa al Corolario 3.15 de [6] que es el que da la solución al problema de optimización, sin usar su teorema principal y otras estimaciones que se presentan en [6].

En el apéndice se encuentran los resultados de un estudio de simulación sobre la distribución empírica espectral de matrices laplacianas con diferentes distribuciones y tamaños de matrices. Estas simulaciones nos llevaron a plantear la conjetura presentada en el Capítulo 2 sobre una aproximación a la densidad de la convolución libre de la distribución semicircular y la distribución normal estándar mediante una mezcla de éstas.

Capítulo 1

Preliminares

Este capítulo contiene los elementos de teoría de gráficas, matrices y optimización necesarios para el desarrollo de esta tesis. La primera sección trata sobre las matrices asociadas a gráficas haciendo énfasis principal en la matriz laplaciana y en cómo este tipo de matrices se puede generalizar al contexto de matrices aleatorias. La segunda sección aborda los conceptos básicos de programación semidefinida necesarios para esta tesis, los cuales fueron tomados de [12]. Por último, en la tercer sección mostramos cómo es la solución de un problema específico de programación semidefinida, el cuál nos servirá para entender las aplicaciones del Capítulo 3.

1.1. Elementos de gráficas y matrices asociadas

En el contexto de gráficas (aleatorias y no aleatorias) es importante estudiar las propiedades espectrales relativas a matrices que se definen a partir de las gráficas, ya que nos proporcionan información acerca de las mismas. Recordemos que una *gráfica* es un par ordenado $G = (V, E)$ compuesto de un conjunto V de vértices y un conjunto $E \subset V \times V$ de aristas, las cuales nos indican que pares

de vértices están conectados. Una gráfica se denomina *simple* si no es posible tener la arista (i, i) , esto es, es una gráfica sin lazos (o loops). Para cada vértice $v \in V$ definimos su *grado*¹ como el número de vértices incidentes a él, es decir, el número de vértices con el que comparte una arista y lo denotamos por $deg(v)$. Denotamos por K_n a la *gráfica completa* de tamaño n , como aquella gráfica con n vértices en las que todos sus grados son iguales a $n - 1$, es decir, cada vértice está conectado con todos los demás vértices.

Se define la *matriz de adyacencia* $A = A_G = (a_{ij})$ de la gráfica simple G como

$$A_{ij} = \begin{cases} 1 & \text{si } (i, j) \in E \\ 0 & \text{en otro caso.} \end{cases}$$

Esto es, la matriz que nos indica cuáles vértices de la gráfica están conectados y cuáles no. A su vez, se define la *matriz laplaciana* $L_G = (L_{ij})$ asociada a la gráfica simple G como

$$L_{ij} = \begin{cases} -a_{ij} & \text{si } i \neq j \\ \sum_{k=1}^n a_{ik} & \text{si } i = j. \end{cases}$$

Es importante notar que los elementos de la diagonal de esta matriz son el grado de cada vértice, además de que ambas matrices son cuadradas de tamaño igual al número de vértices de la gráfica y simétricas. Para el caso de la gráfica completa K_n su matriz de adyacencia y laplaciana están dados por

$$A_{K_n} = \mathbf{1}_n \mathbf{1}_n^t - \mathbf{I}_n, \quad L_{K_n} = (n - 1)\mathbf{I}_n - A_{K_n}$$

donde $\mathbf{1}_n$ es el vector de tamaño n cuyas entradas son iguales a 1 y \mathbf{I}_n es la matriz identidad $n \times n$.

¹Es usual considerar a vértices que comparten una misma arista como vecinos, de esta manera el grado de v es el número de vecinos.

Este tipo de matrices se pueden extender de manera natural a gráficas aleatorias, es decir, gráficas que se construyen de acuerdo a algún modelo aleatorio. Un primer modelo es aquel que asigna a cada par de vértices una arista con una probabilidad fija que puede depender de la cantidad de vértices. Este es el *modelo de Erdős-Rényi* [10, 11] el cual denotaremos por $\mathcal{G}(n, p)$ donde n es el número de vértices y $p = p_n \in [0, 1]$ es la probabilidad de asignar una arista a cada par de vértices.

Para las gráficas de Erdős-Rényi es posible definir matrices de adyacencia y laplacianas de la misma manera que en gráficas deterministas, solo que en este caso estas matrices serán aleatorias. La matriz de adyacencia asociada a una gráfica $G \in \mathcal{G}(n, p)$ no es mas que una matriz aleatoria simétrica cuyas entradas por fuera de la diagonal son variables aleatorias independientes Bernoulli $Ber(p)$ y cuya diagonal es cero, mientras que la matriz laplaciana será una matriz aleatoria simétrica cuyas entradas por fuera de la diagonal son el negativo de una Bernoulli $Ber(p)$ y en la diagonal serán Binomiales $Bin(n - 1, p)$.

Es posible estudiar las matrices laplacinas en un contexto más general, esto es, sin pensar que provienen de una gráfica y considerar las entradas de estas matrices con otra distribución (es decir, diferentes a Bernoulli). Consideremos $\{\xi_{ij}^{(n)}; 1 \leq i < j \leq n, n \geq 2\}$ una sucesión de variables aleatorias, definimos la *matriz de tipo adyacencia* como

$$A = A_n = \begin{pmatrix} 0 & \xi_{12}^{(n)} & \xi_{13}^{(n)} & \cdots & \xi_{1n}^{(n)} \\ \xi_{21}^{(n)} & 0 & \xi_{23}^{(n)} & \cdots & \xi_{2n}^{(n)} \\ \xi_{31}^{(n)} & \xi_{32}^{(n)} & 0 & \cdots & \xi_{3n}^{(n)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \xi_{n1}^{(n)} & \xi_{n2}^{(n)} & \xi_{n3}^{(n)} & \cdots & 0 \end{pmatrix}$$

donde $\xi_{ij} = \xi_{ji}$ para $i, j = 1, \dots, n$. Así mismo se define la *matriz laplaciana*

como

$$L = L_{[A]} = \begin{pmatrix} \sum_{j \neq 1} \xi_{1j}^{(n)} & -\xi_{12}^{(n)} & -\xi_{13}^{(n)} & \cdots & -\xi_{1n}^{(n)} \\ -\xi_{21}^{(n)} & \sum_{j \neq 2} \xi_{2j}^{(n)} & -\xi_{23}^{(n)} & \cdots & -\xi_{2n}^{(n)} \\ -\xi_{31}^{(n)} & -\xi_{32}^{(n)} & \sum_{j \neq 3} \xi_{3j}^{(n)} & \cdots & -\xi_{3n}^{(n)} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ -\xi_{n1}^{(n)} & -\xi_{n2}^{(n)} & -\xi_{n3}^{(n)} & \cdots & \sum_{j \neq n} \xi_{nj}^{(n)} \end{pmatrix}.$$

De esta manera obtenemos que toda matriz laplaciana está definida a partir de una matriz simétrica. En el caso de gráficas esta matriz tiene diagonal cero, puesto que estamos considerando gráficas simples, pero para el estudio del laplaciano podemos considerar que vienen de matrices simétricas sin tener diagonal cero. De esta manera se tiene la siguiente definición de matriz laplaciana (general):

Definición 1.1. Sea $X = (X_{ij})$ una matriz $n \times n$ simétrica. Definimos la matriz laplaciana L_X de X como

$$L_X = D_X - X$$

donde D_X es la matriz diagonal cuyas entradas están dadas por

$$(D_X)_{ii} = \sum_{j=1}^n X_{ij}.$$

Con esta definición tenemos que para la gráfica completa K_n su matriz laplaciana está dada por L_{11^t} .

Para una matriz $n \times m$ A se define la *norma espectral* $\|A\|$ como la raíz cuadrada del mayor eigenvalor $\lambda_{\max}(A^*A)$, donde A^* denota la transpuesta conjugada

de la matriz A . Como caso particular, tenemos que si A es una matriz $n \times n$ simétrica

$$\|A\| = \max\{-\lambda_{\min}(A), \lambda_{\max}(A)\}.$$

Es importante notar que $\lambda_{\max}(\cdot)$ no es una norma en general, pero sin embargo cumple la desigualdad triángular como indicamos en el siguiente lema.

Lema 1.1 (Desigualdad triángular). *Para A, B matrices $n \times n$ simétricas tenemos*

$$\lambda_{\max}(A + B) \leq \lambda_{\max}(A) + \lambda_{\max}(B).$$

Demostración. Sean $k_1, k_2 > 0$ suficientemente grandes tales que las matrices $A + k_1\mathbf{I}_n$ y $B + k_2\mathbf{I}_n$ sean definidas no negativas. Así

$$\begin{aligned} \lambda_{\max}(A + k_1\mathbf{I}_n + B + k_2\mathbf{I}_n) &= \|A + k_1\mathbf{I}_n + B + k_2\mathbf{I}_n\| \\ &\leq \|A + k_1\mathbf{I}_n\| + \|B + k_2\mathbf{I}_n\| \\ &= \lambda_{\max}(A + k_1\mathbf{I}_n) + \lambda_{\max}(B + k_2\mathbf{I}_n) \\ &= \lambda_{\max}(A) + k_1 + \lambda_{\max}(B) + k_2. \end{aligned}$$

Como $\lambda_{\max}(A + k_1\mathbf{I}_n + B + k_2\mathbf{I}_n) = \lambda_{\max}(A + B) + k_1 + k_2$, de la desigualdad anterior se sigue el resultado. □

1.2. Propiedades de las matrices laplacianas

La siguiente proposición resume propiedades de matrices laplacianas usadas en esta tesis. Para una matriz A de tamaño n , denotamos sus eigenvalores por $\lambda_1(A) \leq \lambda_2(A) \leq \dots \leq \lambda_n(A) = \lambda_{\max}(A)$. Aún cuando es usual que el mayor

eigenvalor sea denotado por λ_1 , en el contexto de gráficas se toma al revés dado que el menor eigenvalor es siempre igual a cero para gráficas deterministas.

Proposición 1.1. *Sea L una matriz $n \times n$ laplaciana aleatoria. Entonces*

- (a) *0 es un eigenvalor de L ,*
- (b) *$L_{[X]}$ es semidefinida no negativa si y sólo si las entradas por fuera de la diagonal de $X = (X_{ij})$ son no negativas.*
- (c) *Si L es la matriz laplaciana asociada a una gráfica completa entonces sus eigenvalores son 0 con multiplicidad 1 y n con multiplicidad $n - 1$.*

Demostración. (a) Se sigue del hecho que tanto la suma de las entradas de las columnas como las filas son cero.

(b) Sea $z \in \mathbb{C}^n$ un vector no nulo, y

$$zLz^t = \sum_{j=1}^n z_j \sum_{i=1}^n z_i L_{ij}.$$

Por la estructura de L ,

$$\sum_{i=1}^n z_i L_{ij} = z_j \sum_{i=1}^n X_{ij} - \sum_{i=1}^n z_i X_{ij},$$

luego,

$$\begin{aligned}
zLz^t &= \sum_{j=1}^n z_j \left(z_j \sum_{i=1}^n X_{ij} - \sum_{i=1}^n z_i X_{ij} \right) \\
&= \sum_{j=1}^n z_j^2 \sum_{i=1}^n X_{ij} - \sum_{j=1}^n z_j \sum_{i=1}^n z_i X_{ji} \\
&= \sum_{j=1}^n z_j^2 \sum_{i=1}^n X_{ij} - 2 \sum_{1 \leq i < j \leq n} z_i z_j X_{ij} \\
&= \sum_{1 \leq i < j \leq n} (z_i^2 - 2z_i z_j + z_j^2) X_{ij} \\
&= \sum_{1 \leq i < j \leq n} (z_i - z_j)^2 X_{ij}.
\end{aligned}$$

Por lo tanto, si la matriz $L_{[X]}$ es semidefinida positiva, entonces tomando z el vector con una entrada igual a uno y el resto cero, garantizamos que cada una de las variables aleatorias son no negativas. Por otra parte, si las variables son no negativas, entonces zLz^t es no negativo para cualquier vector $z \in \mathbb{C}^n$.

(c) Recordemos que el laplaciano de una gráfica completa es L_{11^t} y por el inciso (a) sabemos que cero es eigenvalor. Notemos que el rango de esta matriz es $n - 1$, luego por el teorema del rango y nulidad tenemos que la nulidad de L_{11^t} es 1. Como

$$L_{11^t} \mathbf{1}^t = \mathbf{0}$$

entonces $\mathbf{1}$ genera el núcleo de L_{11^t} y por lo tanto 0 es eigenvalor con multiplicidad 1.

Sea $v = (v_1, v_2, \dots, v_n)$ vector ortogonal a $\mathbf{1}$, entonces

$$\begin{aligned} L_{\mathbf{1}\mathbf{1}^t}v &= \left((n-1)v_1 - \sum_{i \neq 1} v_i, (n-1)v_2 - \sum_{i \neq 2} v_i, \dots, (n-1)v_n - \sum_{i \neq n} v_i \right) \\ &= \left(nv_1 - \sum_{i=1}^n v_i, nv_2 - \sum_{i=1}^n v_i, \dots, nv_n - \sum_{i=1}^n v_i \right) \\ &= n(v_1, v_2, \dots, v_n) \\ &= nv. \end{aligned}$$

Por lo tanto n es eigenvalor de $L_{\mathbf{1}\mathbf{1}^t}$ con multiplicidad $n-1$.

□

El siguiente resultado es importante y será utilizado en los siguientes capítulos dado que nos da una manera de relacionar el segundo eigenvalor del negativo del laplaciano más el laplaciano de una gráfica completa con el mayor eigenvalor del primero cuando éste es menor al tamaño de la matriz.

Proposición 1.2. *Sea L una matriz laplaciana $n \times n$, entonces $\lambda_2(L_{\mathbf{1}\mathbf{1}^t} - L) > 0$ si y sólo si $\lambda_{\max}(L) < n$.*

Demostración. Por la Proposición 1.1(c) los eigenvalores de $L_{\mathbf{1}\mathbf{1}^t}$ son 0 con multiplicidad 1 y n con multiplicidad $n-1$. Notemos que si $\alpha > 0$ es eigenvalor de L con eigenvector v ortogonal a $\mathbf{1}$, entonces

$$(L_{\mathbf{1}\mathbf{1}^t} - L)v = L_{\mathbf{1}\mathbf{1}^t}v - Lv = nv - \alpha v = (n - \alpha)v,$$

así $n - \alpha$ es eigenvalor de $L_{\mathbf{1}\mathbf{1}^t} - L$. Es decir, todo ²eigenvalor no nulo de $L_{\mathbf{1}\mathbf{1}^t} - L$ es de la forma $n - \alpha_i$ para algún i donde $\{\alpha_i\}$ son los eigenvalores no nulos de L .

²ésto se tiene por que $\mathbf{1}$ es eigenvector asociado al valor propio cero, así cualquier vector ortogonal a $\mathbf{1}$ no puede tener como eigenvalor al cero.

Notemos además que $L_{11^t} - L$ es matriz laplaciana y 0 es uno de sus eigenvalores. Sean $\{\gamma_1, \gamma_2, \dots, \gamma_n\}$ los eigenvalores de $L_{11^t} - L$. Consideremos $\Lambda_1 = \{\gamma_{(i)}\}$ y $\Lambda_2 = \{\alpha_{(i)}\}$ los eigenvalores ordenados no cero de $L_{11^t} - L$ y L respectivamente, así para todo $j \in \{1, 2, \dots, |\Lambda_1|\}$ tenemos que

$$\gamma_j = n - \alpha_k \quad \text{para algún } k \in \{1, 2, \dots, |\Lambda_2|\}.$$

Notemos que $\gamma_{\min} = n - \alpha_{\max}$.

Luego si $\lambda_2(L_{11^t} - L) > 0$ se tiene que $n - \alpha_{\max} > 0$ lo que implica que $n > \alpha_{\max} = \lambda_{\max}(L)$.

Por otra parte, si $\lambda_{\max}(L) < n$ se tiene que $0 < n - \alpha_{\max} = \lambda_2(L_{11^t} - L)$.

□

1.3. Programación semidefinida

La optimización convexa es una rama de la optimización que se encarga de estudiar problemas del tipo

$$\begin{aligned} \text{mín } & f_0(x) \\ \text{s.t } & f_1(x) \leq 0, \dots, f_N(x) \leq 0 \\ & a'_1 x = b_1, \dots, a'_m x = b_m \end{aligned} \quad (1.3.1)$$

donde $f_i : D \rightarrow \mathbb{R}$ son funciones convexas con $D \subset \mathbb{R}^n$ convexo.

En esta sección resumimos los elementos necesarios de este tema para estudiar las aplicaciones del Capítulo 3. El material se basa en [12], referencia que recomendamos para un estudio sistemático del tema de programación semidefinida.

Decimos que una solución de (1.3.1) es factible si cumple todas las restricciones. De esta manera, podemos definir el conjunto de posibles soluciones (que es un conjunto convexo) como

$$\bigcap_{i=1}^N \{x \in D : f_i(x) \leq 0\} \cap \bigcap_{j=1}^m \{x \in \mathbb{R}^n : a'_j x = b_j\}.$$

Una clase particular de problemas en optimización convexa es el denominado *Programación Cónica* (conic programming). En este tipo de problemas las soluciones se encuentran en un conjunto convexo, más precisamente consideremos $K \subset \mathbb{R}^n$ fijo un conjunto convexo, entonces el problema de programación cónica está dado por

$$\begin{aligned} \text{máx} \quad & c'x \\ \text{s.t} \quad & x \in K \\ & a'_1 x = b_1, \dots, a'_m x = b_m. \end{aligned} \tag{1.3.2}$$

En esta clase de problemas, se encuentran la programación lineal, programación cónica de segundo orden (CQD) y programación semidefinida (SDP por sus siglas en inglés).

La programación semidefinida se encarga de estudiar problemas de optimización convexa en los cuales el conjunto de soluciones K pertenece a las matrices semidefinidas no negativas las cuales denotamos por $X \succeq 0$. También denotamos la traza³ de una matriz X como $Tr(X)$ y definimos $\langle X, Y \rangle = Tr(X^t Y)$ donde X^t e Y son matrices $m \times n$ ambas con entradas reales.

La forma estándar de un problema de programación semidefinida está dado por

$$p^* = \sup\{\langle C, X \rangle : X \succeq 0, \langle A_1, X \rangle = b_1, \dots, \langle A_m, X \rangle = b_m\}. \tag{1.3.3}$$

Dentro de esta teoría es usual hablar del problema dual. Definimos el cono dual K^* del cono K (i.e., es cerrado bajo sumas y multiplicación por escalares) como

$$K^* = \{y \in \mathbb{R}^n : y'x \geq 0, \forall x \in K\}.$$

³ no normalizada.

Con la anterior, definimos el problema de programación semidefinida dual como

$$d^* = \inf \left\{ \sum_{j=1}^m y_j b_j : y_1, \dots, y_m \in \mathbb{R}^m, \sum_{j=1}^m y_j A_j - C \succeq 0 \right\}. \quad (1.3.4)$$

La ventaja de trabajar con el problema dual es que al tener una solución factible para este problema, se tiene que es automáticamente una solución para el problema estándar. Además, en muchos casos el problema dual es mucho más sencillo de solucionar.

Teorema 1.1 (*Dualidad Débil*). *Sea x una solución factible del problema de programación semidefinida estándar (1.3.3) e y una solución factible del problema dual (1.3.4). Entonces*

$$c'x \leq b'y$$

en particular, $p^* \leq d^*$.

Demostración. Notemos que

$$\sum_{j=1}^m y_j b_j = \sum_{j=1}^m y_j (a'_j x) = \left(\sum_{j=1}^m y_j a_j \right)^T x \geq c'x$$

donde la última desigualdad se da por $\sum_{j=1}^m y_j a_j - c \in K^*$.

□

1.4. Programación semidefinida en $\{-1, 1\}^n$

El siguiente problema de programación semidefinida aparece en dos problemas de sincronización y en modelos de bloques estocásticos que se desarrollan en el Capítulo 3.

Consideremos el siguiente problema de SDP: Dada Y una matriz de observaciones, el objetivo es encontrar

$$\begin{aligned} \text{máx} \quad & \text{Tr}(YX) \\ \text{s.t} \quad & X_{ii} = 1 \\ & X \succeq 0. \end{aligned} \tag{1.4.1}$$

Este es un problema de programación semidefinida. Para ver esto tomemos A_i como la matriz tal que es 1 en la entrada (i, i) y cero en el resto, así el problema anterior se puede escribir como (1.3.3) de la siguiente manera:

$$\text{máx}\{\langle X, Y \rangle : X \succeq 0, \langle X, A_i \rangle = 1 \text{ para } i = 1, \dots, n\}. \tag{1.4.2}$$

Ahora, como los b_j de la definición del problema de programación semidefinida en este caso son iguales a 1, tenemos $\sum_{j=1}^n y_j b_j = \sum_{j=1}^n y_j = \text{Tr}(D)$ donde D es una matriz diagonal. De esta manera tenemos que el problema dual (1.3.4) está dado por

$$\text{mín}\{\text{Tr}(D) : D \text{ Diagonal}, D - Y \succeq 0\}. \tag{1.4.3}$$

Usando el concepto de dualidad débil (Teorema 1.1), si X y D son soluciones factibles del problema (1.4.2), (1.4.3) respectivamente, se tiene que

$$\text{Tr}(YX) \leq \text{Tr}(D).$$

Así, como X y $D - Y$ son semidefinidas no negativas, se tiene que

$$0 \leq \text{Tr}((D - Y)X) = \text{Tr}(DX) - \text{Tr}(YX) = \text{Tr}(D) - \text{Tr}(YX).$$

Por lo tanto, si encontramos una matriz D diagonal que solucione el problema dual tal que $\text{Tr}(D) = \text{Tr}(Yxx^t)$, entonces xx^t es solución óptima del problema

original gracias al teorema de dualidad débil. Para garantizar unicidad es necesario además que $\lambda_2(D - Y) > 0$ [1, 6].

Con todo lo anterior obtenemos el siguiente resultado que se conoce como certificado dual (dual certificate)

Lema 1.2 (Certificado Dual). *Sea Y una matriz $n \times n$ semidefinida no negativa y $x \in \{-1, 1\}^n$. Si existe una matriz diagonal D tal que $\text{Tr}(D) = x^t Y x$, $D - Y \succeq 0$ y $\lambda_2(D - Y) > 0$, entonces $X = xx^t$ es la única solución óptima del problema (1.4.1).*

Consideremos el siguiente candidato para el certificado dual

$$D_{ii} = \sum_{j=1}^n Y_{ij} x_i x_j.$$

Notemos que, usando la Definición 1.1, podemos escribir esta matriz como

$$D = D_{[\text{diag}(x)Y\text{diag}(x)]}$$

y

$$\text{Tr}(D) = \sum_{i=1}^n D_{ii} = \sum_{i=1}^n \sum_{j=1}^n Y_{ij} x_i x_j = x^t Y x,$$

además $(D - Y)x = 0$. La matriz D cumple con las hipótesis del lema anterior, con lo cual obtenemos el siguiente resultado que usaremos en el capítulo 3.

Proposición 1.3 (Certificado Dual). *Sea Y una matriz $n \times n$ semidefinida no negativa y $x \in \{-1, 1\}^n$. Sea $D = D_{[\text{diag}(x)Y\text{diag}(x)]}$. Si $\lambda_2(D - Y) > 0$, entonces $X = xx^t$ es la única solución óptima del problema de programación semidefinida (1.4.1).*

Capítulo 2

Sobre el espectro asintótico del laplaciano

Este capítulo presenta resultados asintóticos sobre la distribución empírica espectral y el mayor eigenvalor de una matriz laplaciana aleatoria. Comenzamos mencionando los resultados de Ding-Jiang [9] que están basados en el artículo pionero de Bryc-Dembo-Jiang [8]. Usando estos resultados proponemos un enfoque distinto al estudio de estimar el mayor eigenvalor de una matriz laplaciana que hace Bandeira en [6], lo cual nos permite obtener extensiones que dan un enfoque unificado a las aplicaciones presentadas en el Capítulo 3. Así mismo se presentan consecuencias que obtuvimos del resultado de [9] sobre la convergencia del mayor eigenvalor en matrices laplacianas a bloques. Presentamos un resultado que obtuvimos sobre la distribución asintótica del mayor eigenvalor de una matriz laplaciana con entradas normales estándar.

2.1. Distribución empírica espectral

La distribución empírica espectral juega un papel muy importante en la teoría de matrices aleatorias. En su trabajo pionero, Wigner [14] demostró que para una *matriz de Wigner* W , la distribución empírica espectral de $\frac{1}{\sqrt{n}}W$ converge débilmente casi seguramente a la distribución del semicírculo S en $[-2, 2]$

$$S(dx) = \frac{1}{2\pi} \sqrt{4 - x^2} \mathbb{1}_{|x| \leq 2} dx.$$

Por matriz de Wigner $W = (W_{ij})$ nos referimos a una matriz simétrica $n \times n$ cuyas entradas $\{W_{ij}, 1 \leq i < j \leq n\}$ son variables aleatorias independientes con la misma distribución de media cero y varianza uno mientras que $\{W_{ii}, 1 \leq i \leq n\}$ son variables aleatorias independientes de media cero y varianza dos independientes de las entradas fuera de la diagonal.

Una pregunta natural es si es posible obtener resultados asintóticos similares para matrices aleatorias de adyacencia y laplacianas generales.

En el caso de matrices de adyacencia, la dificultad radica en que los elementos de la diagonal son cero por lo que no tienen varianza positiva. En el caso de las matrices laplacianas, el reto es que los elementos de la diagonal no son independientes. A continuación se presentan los resultados de [9] que abordan los problemas arriba mencionados de manera general, en el sentido de que las variables aleatorias consideradas siguen siendo independientes pero no necesariamente con la misma distribución.

En el resto del capítulo y la tesis se tiene la siguiente suposición sobre las variables aleatorias.

Condición 1. Sea $X_n = (\xi_{ij}^{(n)})$ una matriz simétrica $n \times n$ con $\{\xi_{ij}^{(n)}; 1 \leq i < j \leq n, n \geq 2\}$ variables aleatorias definidas en el mismo espacio de probabilidad y $\{\xi_{ij}^{(n)}; 1 \leq i < j \leq n\}$ independientes para cada $n \geq 2$ (no necesariamente con la

misma distribución) con $\xi_{ij}^{(n)} = \xi_{ji}^{(n)}$, $E(\xi_{ij}^{(n)}) = \mu_n$, $\text{Var}(\xi_{ij}^{(n)}) = \sigma_n^2 > 0$ para todo $1 \leq i < j \leq n$ y $n \geq 2$ y

$$\sup_{1 \leq i < j \leq n, n \geq 2} E|(\xi_{ij}^{(n)} - \mu_n)/\sigma_n|^p < \infty$$

para algún $p > 0$.

Para el caso de matrices aleatorias de tipo adyacencia se tiene el siguiente resultado (Teorema 3 de [9]).

Teorema 2.1. *Sea $A_n = (\xi_{ij}^{(n)})$ una matriz de tipo adyacencia y consideremos $\omega_{ij}^{(n)} := (\xi_{ij}^{(n)} - \mu_n)/\sigma_n$ para todo i, j, n . Supongamos que la Condición 1 se satisface con $p = 2$ y*

$$\max_{1 \leq i < j \leq n} E\{(\omega_{ij}^{(n)})^2 \mathbb{I}(|\omega_{ij}^{(n)}| \geq \varepsilon\sqrt{n})\} \rightarrow 0$$

cuando $n \rightarrow \infty$ para todo $\varepsilon > 0$. Sea

$$\tilde{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}\left\{\frac{\lambda_i(A_n) + \mu_n}{\sqrt{n}\sigma_n} \leq x\right\}, \quad x \in \mathbb{R}.$$

Entonces, casi seguramente, \tilde{F}_n converge débilmente a la ley del semicírculo S en $[-2, 2]$.

Para el caso de la matriz laplaciana se tiene lo siguiente (Teorema 2 de [9])

Teorema 2.2. *Sea $L_n = L_{[X_n]}$ la matriz laplaciana de $X_n = (\xi_{ij}^{(n)})$. Supongamos que las entradas de X_n satisfacen la Condición 1 para algún $p > 4$. Sea*

$$\tilde{F}_n(x) = \frac{1}{n} \sum_{i=1}^n \mathbb{I}\left\{\frac{\lambda_i(L_n) - n\mu_n}{\sqrt{n}\sigma_n} \leq x\right\}, \quad x \in \mathbb{R}.$$

Entonces, cuando $n \rightarrow \infty$, con probabilidad 1, \tilde{F}_n converge débilmente a la convolución libre $\gamma_M = S \boxplus \gamma$ de la ley del semicírculo S en $[-2, 2]$ y la distribución normal estándar γ . La medida γ_M es simétrica, no aleatoria con soporte no acotado y con densidad acotada, que no depende de la distribución de $\{\xi_{ij}^{(n)}; 1 \leq i < j \leq n, n \geq 2\}$.

Observación. En [8] se dan resultados para la distribución asintótica empírica espectral de matrices simétricas del tipo Hankel, Markov y Toeplitz. Extensiones del resultado anterior al caso de laplaciano de modelos de bloques estocásticos fueron obtenidos en 2015 en [2].

2.2. Densidad de la distribución asintótica espectral

Realizamos un estudio de simulación para obtener intuición sobre la distribución de la convolución libre γ_M de la ley del semicírculo y la distribución normal estándar. Simulamos matrices laplacianas de diferentes tamaños con entradas por fuera de la diagonal de media cero y varianza uno, con el fin de ver cómo es su distribución empírica espectral.

Después de realizar numerosas simulaciones concluimos que una mezcla de la densidad del semicírculo S en $[-2, 2]$ y normal estándar γ parece ajustar la densidad como lo soporta la Figura 2.2 y otras que se presentan en el apéndice.

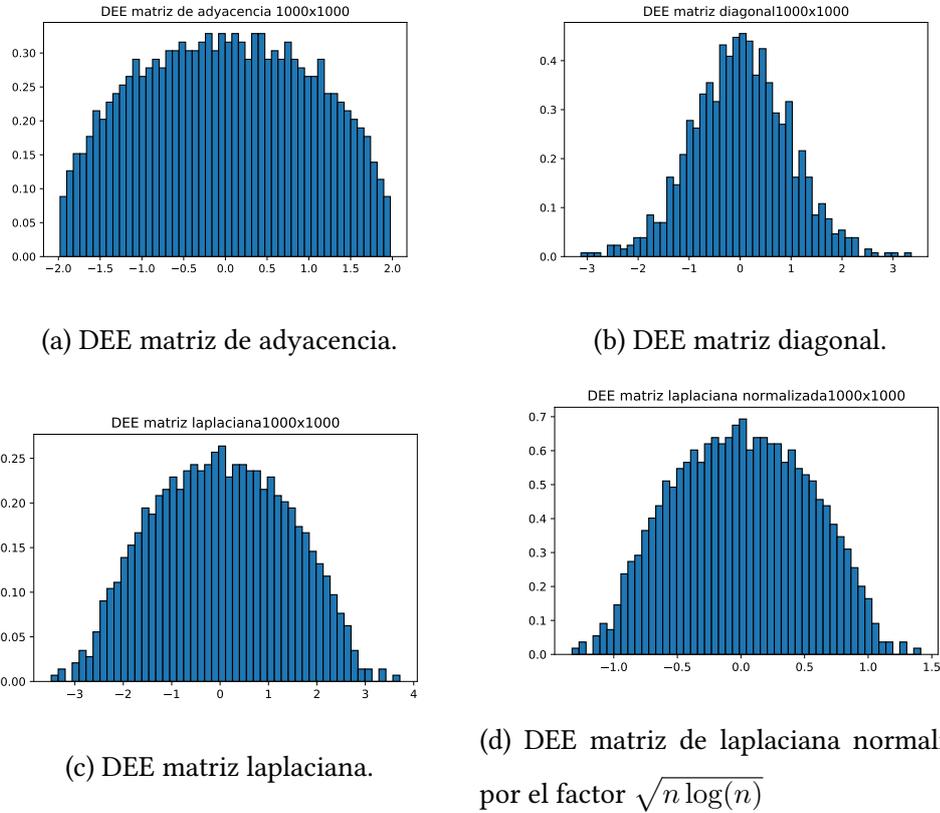


Figura 2.1: Histogramas de la distribución empírica espectral (DEE) de la matriz de adyacencia con entradas normales con media cero y varianza 1 de tamaño 1000 y la diagonal y laplaciana dadas por las entradas de la misma. En la Figura 2.1a vemos como se aproxima a la densidad del semicírculo S en $[-2, 2]$ como lo indica el Teorema 2.1. En la Figura 2.1b se ve el comportamiento de los elementos de la diagonal, que al ser sumas estandarizadas de variables aleatorias independientes se espera que tengan un comportamiento gaussiano. En las Figuras 2.1c, 2.1d en virtud del Teorema 2.2 tenemos una aproximación a la densidad de la convolución libre γ_M , al normalizar la matriz laplaciana por $\sqrt{n \log n}$ notamos que el mayor eigenvalor se está acercado a 1.5 que es cercano al valor $\sqrt{2}$, que es el límite en probabilidad de $\lambda_{\max}(A)/\sqrt{n \log(n)}$ como se presenta en la siguiente sección.

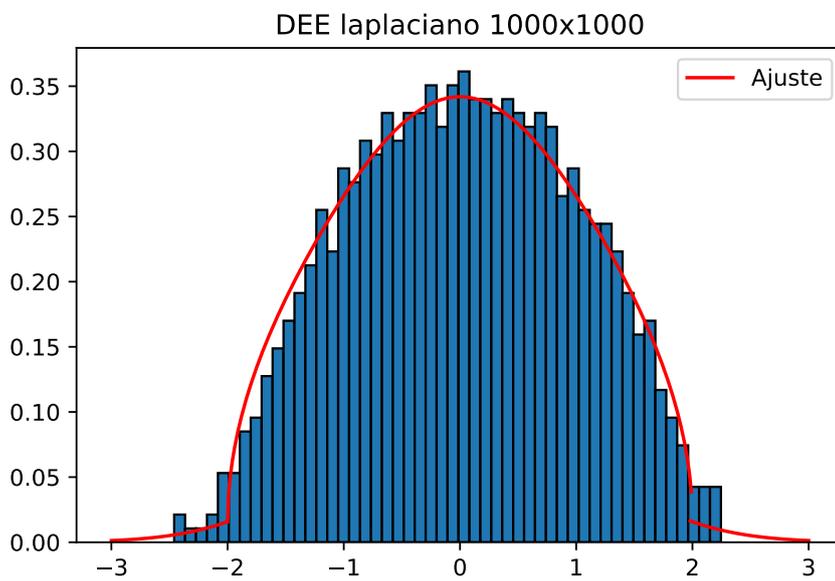


Figura 2.2: Ajuste al histograma de la distribución empírica espectral de una matriz laplaciana de tamaño 1000 normalizada por $\sqrt{2 * 1000}$

En base a lo anterior formulamos la siguiente conjetura.

Conjetura 2.1. *La densidad de la convolución libre $\gamma_M = S \boxplus \gamma$ es o se puede aproximar por la mezcla de las densidades semicircular y normal. Más aún, si la matriz laplaciana viene de una matriz cuyas entradas por fuera de la diagonal tienen media cero y varianza σ^2 entonces*

$$f_{S \boxplus \gamma}(x) = \frac{\alpha}{2\sigma\sqrt{2}} f_S\left(\frac{x}{\sigma\sqrt{2}}\right) + \frac{1-\alpha}{2\sigma\sqrt{2}} f_\gamma\left(\frac{x}{\sigma\sqrt{2}}\right) \quad (2.2.1)$$

$$= \frac{\alpha}{2\sigma\sqrt{2}} f_{S_\sigma}(x) + \frac{1-\alpha}{2\sigma\sqrt{2}} f_{\gamma_\sigma}(x) \quad (2.2.2)$$

donde S_σ es la distribución del semicírculo en $[-\sqrt{2}\sigma, \sqrt{2}\sigma]$ y γ_σ es la distribución normal de media cero y varianza $2\sigma^2$ y $\alpha = \sqrt{2}/2$.

El peso α parece representar el cociente entre los límites inferior y superior casi seguros \liminf y \limsup del mayor eigenvalor de la matriz laplaciana normalizada $\lambda_{\max}(L)/\sqrt{n \log n}$. También parece ser el cociente entre el límite $\lambda_{\max} L/\sqrt{n \log(n)}$ y el límite del mayor eigenvalor de una matriz de Wigner (ambos resultados se presentan en la siguiente sección).

Como caso particular tenemos

$$\frac{f_{S \boxplus \gamma}(x)}{\sqrt{2}} = \alpha f_S(x) + (1-\alpha) f_\gamma(x). \quad (2.2.3)$$

Otro enfoque para encontrar la densidad de la distribución asintótica es considerando la distribución empírica espectral de matrices a bloques, lo cual presentamos de forma detallada como sigue. La razón de considerar esto es porque mezclas de distribuciones como las anteriores, ocurren de manera natural como distribución empírica espectral de matrices diagonales a bloques.

Teorema 2.3. *Sea*

$$X = \left[\begin{array}{c|c} A & 0 \\ \hline 0 & B \end{array} \right]$$

con A y B matrices de tamaño $n \times n$ y $m \times m$ respectivamente. Entonces la distribución empírica espectral de X es una combinación convexa de las distribuciones empírica espectral de A y B , de forma más específica,

$$F^X(x) = \frac{n}{n+m} F^A \left(\sqrt{\frac{n+m}{n}} x \right) + \frac{m}{n+m} F^B \left(\sqrt{\frac{n+m}{m}} x \right). \quad (2.2.4)$$

Demostración. Notemos que X es una matriz de tamaño $(n+m) \times (n+m)$, además por su estructura sus eigenvalores λ_i serán la unión de los eigenvalores de A y B . Con lo anterior, tenemos que su distribución empírica espectral está dada por

$$\begin{aligned} F^X(x) &= \frac{1}{n+m} \sum_{i=1}^{n+m} \delta_{\lambda_i(X)/\sqrt{n+m}}(x) \\ &= \frac{1}{n+m} \left(\sum_{i=1}^n \delta_{\lambda_i(A)/\sqrt{n+m}}(x) + \sum_{j=1}^m \delta_{\lambda_j(B)/\sqrt{n+m}}(x) \right). \end{aligned}$$

Pero

$$\begin{aligned} \delta_{\lambda_i(A)/\sqrt{n+m}}(x) &= \delta_{\lambda_i(A)/\sqrt{n}} \left(\sqrt{\frac{n+m}{n}} x \right) \\ \delta_{\lambda_j(B)/\sqrt{n+m}}(x) &= \delta_{\lambda_j(B)/\sqrt{m}} \left(\sqrt{\frac{n+m}{m}} x \right), \end{aligned}$$

así

$$\begin{aligned}
F^X(x) &= \frac{1}{n+m} \left(\sum_{i=1}^n \delta_{\lambda_i(A)/\sqrt{n}} \left(\sqrt{\frac{n+m}{n}} x \right) + \sum_{j=1}^m \delta_{\lambda_j(A)/\sqrt{m}} \left(\sqrt{\frac{n+m}{m}} x \right) \right) \\
&= \frac{n}{n+m} \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i(A)/\sqrt{n}} \left(\sqrt{\frac{n+m}{n}} x \right) + \frac{m}{n+m} \frac{1}{m} \sum_{j=1}^m \delta_{\lambda_j(A)/\sqrt{m}} \left(\sqrt{\frac{n+m}{m}} x \right) \\
&= \frac{n}{n+m} F^A \left(\sqrt{\frac{n+m}{n}} x \right) + \frac{m}{n+m} F^B \left(\sqrt{\frac{n+m}{m}} x \right)
\end{aligned}$$

□

Como caso particular, cuando $n = m$ tenemos

$$F^X(x) = \frac{1}{2} F^A(\sqrt{2}x) + \frac{1}{2} F^B(\sqrt{2}x). \quad (2.2.5)$$

Recordemos que de la Definición 1.1 de matriz laplaciana tenemos

$$L_X = D_X - X$$

donde D_X es una matriz diagonal y X una matriz simétrica. Ahora, notemos que

$$\left[\begin{array}{c|c} I & I \\ \hline 0 & 0 \end{array} \right] \left[\begin{array}{c|c} D_X & 0 \\ \hline 0 & -X \end{array} \right] \left[\begin{array}{c|c} I & 0 \\ \hline I & 0 \end{array} \right] = \left[\begin{array}{c|c} D_X - X & 0 \\ \hline 0 & 0 \end{array} \right] = \left[\begin{array}{c|c} L_X & 0 \\ \hline 0 & 0 \end{array} \right]. \quad (2.2.6)$$

De esta manera, la representación anterior sugiere que la distribución empírica espectral del laplaciano está dada por una perturbación de la distribución de la matriz a bloques dados por D_X y X , que tiene por distribución empírica espectral a (2.2.5). Esto es un tema de estudio a futuro.

2.3. Mayor eigenvalor

Sea $\lambda_{\max}(W)$ el mayor eigenvalor de una matriz de Wigner W . Bai y Yin [4] demostraron que

$$\frac{1}{\sqrt{n}} \lambda_{\max}(W) \rightarrow 2 \quad c.s. \quad \text{cuando } n \rightarrow \infty. \quad (2.3.1)$$

En virtud del Teorema 2.1 es natural pensar que se tenga el Lema 2.1 si W es una matriz de adyacencia. Sin embargo en el caso de las matrices laplacianas la distribución empírica espectral asintótica es de soporte no acotado y se tiene un resultado diferente (Teorema 2.4 abajo).

El siguiente resultado (Lema 2.1 en [9]) da el comportamiento del mayor eigenvalor para matrices similares a las de Wigner pero con varianza uno en la diagonal. La segunda parte da el comportamiento del mayor eigenvalor para cierto tipo de matrices de tipo adyacencia. Estos resultados serán usados en el Capítulo 3.

Lema 2.1. *Sea $\mathbf{U}_n = (u_{ij}^{(n)})$ una matriz aleatoria simétrica $n \times n$ donde las variables aleatorias $\{u_{ij}^{(n)} : 1 \leq i \leq j \leq n, n \geq 1\}$ están definidas en el mismo espacio de probabilidad. Supongamos que para cada $n \geq 1$, $\{u_{ij}^{(n)} : 1 \leq i \leq j \leq n\}$ son independientes con $\mathbb{E}u_{ij}^{(n)} = 0$, $\text{Var}(u_{ij}^{(n)}) = 1$ para todo $1 \leq i, j \leq n$, y $\sup_{1 \leq i, j \leq n, n \geq 1} \mathbb{E}|u_{ij}^{(n)}|^{6+\delta} < \infty$ para algún $\delta > 0$. Entonces:*

(i)

$$\lim_{n \rightarrow \infty} \frac{\lambda_{\max}(\mathbf{U}_n)}{\sqrt{n}} = 2 \quad c.s. \quad y \quad \lim_{n \rightarrow \infty} \frac{\|\mathbf{U}_n\|}{\sqrt{n}} = 2 \quad c.s.$$

(ii) *Los resultados anteriores se siguen cumpliendo si \mathbf{U}_n es reemplazada por $\mathbf{U}_n - \text{diag}(u_{ij}^{(n)})$.*

Una descripción amplia del comportamiento asintótico del mayor eigenvalor de laplacianos se da en el siguiente teorema que es uno de los dos resultados principales en [9].

Teorema 2.4. *Sea $L^{(n)} = L_{[X_n]}$ la matriz laplaciana de una matriz $n \times n$ simétrica X_n con entradas que satisfacen la Condición 1 para algún $p > 6$. Si $\mu_n = 0$ y $\sigma_n = 1$ para todo $n \geq 2$, entonces:*

(a)

$$\frac{\lambda_{\text{máx}}(L^{(n)})}{\sqrt{n \log n}} \rightarrow \sqrt{2}$$

en probabilidad cuando $n \rightarrow \infty$.

Además, si $\{L^{(2)}, L^{(3)}, \dots\}$ es una sucesión de matrices laplacianas independientes, entonces:

(b)

$$\liminf_{n \rightarrow \infty} \frac{\lambda_{\text{máx}}(L^{(n)})}{\sqrt{n \log n}} = \sqrt{2} \quad \text{c.s.}$$

y

$$\limsup_{n \rightarrow \infty} \frac{\lambda_{\text{máx}}(L^{(n)})}{\sqrt{n \log n}} = 2 \quad \text{c.s.},$$

y la sucesión $\{\lambda_{\text{máx}}(L^{(n)})/\sqrt{n \log n}; n \geq 2\}$ es densa en $[\sqrt{2}, 2]$ c.s.;

(c) las conclusiones en (a), (b) se conservan si $\lambda_{\text{max}}(L^{(n)})$ es remplazado por $\|L_n\|$.

Podemos extender el Teorema 2.4 a ciertas matrices diagonales a bloques

$$B = \begin{bmatrix} B_1 & 0 & \dots & 0 \\ 0 & B_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & B_k \end{bmatrix},$$

donde $\{B_1, B_2, \dots, B_k\}$ son matrices cuadradas del mismo tamaño.

Una consecuencia muy útil del teorema anterior es lo siguiente.

Diremos que una condición A_n se cumple *con alta probabilidad* si para todo $\eta > 0$, existe $n_0 \in \mathbb{N}$ tal que

$$\mathbb{P}(A_n) \geq 1 - \eta \quad \text{para } n > n_0.$$

Corolario 2.1. *Sea $L = L_{[X]}$ una matriz laplaciana de una matriz $n \times n$ simétrica X con entradas que satisfacen la Condición 1 para algún $p > 6$. Si $\mu_n = 0$ y $\sigma_n = 1$ para todo $n \geq 2$, entonces para todo $\epsilon > 0$ con alta probabilidad*

$$\lambda_{\max}(L) \leq \sqrt{(2 + \epsilon)n \log n} \quad (2.3.2)$$

y

$$\lambda_{\max}(L) \geq (2\sqrt{2} - \sqrt{2 + \epsilon})\sqrt{n \log n}. \quad (2.3.3)$$

Demostración. La prueba se sigue del Teorema 2.4 que asegura que

$$\mathbb{P} \left(\left| \frac{\lambda_{\max}(L)}{\sqrt{n \log n}} - \sqrt{2} \right| > \delta \right) \rightarrow 0 \quad \text{cuando } n \rightarrow \infty$$

tomando $\delta = \sqrt{2 + \epsilon} - \sqrt{2} > 0$. \square

Observamos que el lado derecho de (2.3.3) es positivo si $0 < \epsilon < 6$.

Otra consecuencia aún más útil son los siguientes resultados que obtuvimos para el caso de matrices laplacianas a bloques, el cual permite extender modelos de bloques considerados en el Capítulo 3.

Proposición 2.1. *Sea L una matriz laplaciana diagonal a k bloques cada uno de tamaño n/k cuyas entradas por fuera de la diagonal en cada bloque son independientes. Supongamos que la Condición 1 se satisface para algún $p > 6$. Si $\mathbb{E}[L_{ij}] = 0$ y $\text{Var}(L_{ij}) = 1$ para todo $n \geq 2$. Entonces*

$$\frac{\lambda_{\max}(L)}{\sqrt{n \log n}} \rightarrow \sqrt{\frac{2}{k}} \quad (2.3.4)$$

en probabilidad cuando $n \rightarrow \infty$.

Demostración. Sean L_1, L_2, \dots, L_k los k bloques de la matriz L . Para $i = 1, \dots, k$ sea

$$\Omega_i^{(n)} = \{\omega \in \Omega : \lambda_{\max}(L_j) \leq \lambda_{\max}(L_i) \text{ para todo } j\}.$$

Tenemos que

$$\begin{aligned} \mathbb{P} \left(\left| \frac{\lambda_{\max}(L)}{\sqrt{n \log n}} - \sqrt{\frac{2}{k}} \right| > \epsilon \right) &= \sum_{i=1}^k \mathbb{P} \left(\left\{ \left| \frac{\lambda_{\max}(L)}{\sqrt{n \log n}} - \sqrt{\frac{2}{k}} \right| > \epsilon \right\} \cap \Omega_i^{(n)} \right) \\ &\leq \sum_{i=1}^k \mathbb{P} \left(\left| \frac{\lambda_{\max}(L_i)}{\sqrt{n \log n}} - \sqrt{\frac{2}{k}} \right| > \epsilon \right) \\ &\leq \sum_{i=1}^k \mathbb{P} \left(\left| \frac{\lambda_{\max}(L_i)}{\sqrt{\frac{n}{k} \log \frac{n}{k}}} \frac{\sqrt{\frac{n}{k} \log \frac{n}{k}}}{\sqrt{n \log n}} - \sqrt{\frac{2}{k}} \right| > \epsilon \right) \end{aligned}$$

Notemos que cada matriz L_i es de tamaño $n/k \times n/k$ y cumple las hipótesis del Teorema 2.4, así para $i = 1, \dots, k$,

$$\frac{\lambda_{\max}(L_i)}{\sqrt{\frac{n}{k} \log \frac{n}{k}}} \rightarrow \sqrt{2}$$

en probabilidad cuando $n \rightarrow \infty$ y

$$\frac{\sqrt{\frac{n}{k} \log \frac{n}{k}}}{\sqrt{n \log n}} \rightarrow \frac{1}{\sqrt{k}} \quad \text{cuando} \quad n \rightarrow \infty.$$

Luego, una aplicación del teorema de Slutsky nos da que para todo $i = 1, \dots, k$,

$$\frac{\lambda_{\max}(L_i)}{\sqrt{\frac{n}{k} \log \frac{n}{k}}} \frac{\sqrt{\frac{n}{k} \log \frac{n}{k}}}{\sqrt{n \log n}} \rightarrow \sqrt{\frac{2}{k}}$$

en probabilidad cuando $n \rightarrow \infty$. Lo anterior garantiza que para todo $\epsilon > 0$ se tiene que

$$\mathbb{P} \left(\left| \frac{\lambda_{\max}(L)}{\sqrt{n \log n}} - \sqrt{\frac{2}{k}} \right| > \epsilon \right) \rightarrow 0,$$

de donde se obtiene el resultado. \square

De la proposición anterior y siguiendo las ideas del Corolario 2.1 obtenemos la siguiente proposición que nos dan cotas para el mayor eigenvalor de matrices laplacianas a bloques.

Teorema 2.5. *Sea L una matriz laplaciana diagonal a k bloques cada uno de tamaño n/k cuyas entradas por fuera de la diagonal en cada bloque son independientes. Supongamos que la Condición 1 se satisface para algún $p > 6$. Si $\mathbb{E}[L_{ij}] = 0$ y $\text{Var}(L_{ij}) = \sigma^2$ para todo $n \geq 2$. Entonces para todo $\epsilon > 0$ con alta probabilidad*

$$\lambda_{\max}(L) \leq \sigma \sqrt{\left(\frac{2}{k} + \epsilon\right) n \log n} \quad (2.3.5)$$

y

$$\lambda_{\max}(L) \geq \sigma \left(2\sqrt{2} - \sqrt{\frac{2}{k} + \epsilon}\right) \sqrt{n \log n}. \quad (2.3.6)$$

Demostración. La prueba se sigue de la Proposición 2.1 que asegura que

$$\mathbb{P}\left(\left|\frac{\lambda_{\max}(L)}{\sqrt{n \log n}} - \sqrt{\frac{2}{k}}\right| > \delta\right) \rightarrow 0 \quad \text{cuando} \quad n \rightarrow \infty$$

tomando

$$\delta = \sqrt{\frac{2}{k} + \epsilon} - \sqrt{\frac{2}{k}} > 0$$

□

2.4. Estimación del mayor eigenvalor

Otra consecuencia importante del Teorema 2.4 es la estimación del mayor eigenvalor con el mayor elemento de la diagonal del laplaciano. Esto es útil para reducir el costo computacional del cálculo de los eigenvalores.

En esta dirección obtuvimos el siguiente resultando que nos permite proponer un enfoque alternativo y unificado a aplicaciones consideradas en [6] y dar pruebas alternativas a sus resultados; como se muestra en el Capítulo 3.

Teorema 2.6. *Sea L una matriz laplaciana cuyas entradas por fuera de la diagonal son independientes con media cero y varianza σ^2 (no dependiendo de n) cumpliendo*

la Condición 1. Si existen $c_1 > 0$ tal que

$$c_1 \sqrt{\frac{\log n}{n-1}} \leq \sigma \quad (2.4.1)$$

y $c_2 > 0$ tal que con alta probabilidad

$$\sigma \sqrt{(n-1) \log n} \leq c_2 \max_i L_{ii}, \quad (2.4.2)$$

entonces para todo $\epsilon > 0$ y

$$C = c_2 \sqrt{2 + \epsilon} > 0, \quad (2.4.3)$$

se cumple

$$\lambda_{\max}(L) \leq C \left(1 + \frac{\sigma}{c_1 \sqrt{\log n}} \right) \max_i L_{ii} \quad (2.4.4)$$

con alta probabilidad.

Demostración. Consideremos la matriz \hat{L} dada por

$$\hat{L} = \frac{1}{\sigma} L.$$

Notemos que para todo $1 \leq i < j \leq n$ se cumple $\mathbb{E} \hat{L}_{ij} = 0$ y $\mathbb{E} \hat{L}_{ij}^2 = 1$. De esta manera la matriz \hat{L} cumple las hipótesis del Corolario 2.1. Así para todo $\epsilon > 0$,

$$\lambda_{\max}(\hat{L}) \leq \sqrt{(2 + \epsilon)n \log n}$$

y

$$\lambda_{\max}(L) \leq \sigma \sqrt{(2 + \epsilon)n \log n}$$

con alta probabilidad. Ahora notemos que

$$\begin{aligned} \lambda_{\max}(L) &\leq \sqrt{(2 + \epsilon)\sigma^2 n \log n} \\ &\leq \sqrt{(2 + \epsilon)\sigma^2 (n-1) \log n} + \sqrt{(2 + \epsilon)\sigma^2 \log n} \\ &\leq \sqrt{2 + \epsilon} c_2 \max_i L_{ii} + \sigma \sqrt{(2 + \epsilon) \log n} \end{aligned}$$

de donde obtenemos

$$\lambda_{\max}(L) \leq \left(\sqrt{2 + \epsilon} c_2 + \frac{1}{\max_i L_{ii}} \sigma \sqrt{(2 + \epsilon) \log n} \right) \max_i L_{ii}. \quad (2.4.5)$$

Como se cumple

$$c_2 \max_i L_{ii} \geq \sigma \sqrt{(n-1) \log n} \Leftrightarrow \frac{1}{\max_i L_{ii}} \leq \frac{c_2}{\sigma \sqrt{(n-1) \log n}}$$

con alta probabilidad, entonces

$$\frac{1}{\max_i L_{ii}} \sigma \sqrt{(2 + \epsilon) \log n} \leq \frac{c_2 \sigma \sqrt{(2 + \epsilon) \log n}}{\sigma \sqrt{(n-1) \log n}} = \frac{c_2 \sigma \sqrt{(2 + \epsilon)}}{\sigma \sqrt{n-1}}$$

y como

$$\sigma > c_1 \sqrt{\frac{\log n}{n-1}} \Leftrightarrow \frac{1}{\sigma} < \frac{\sqrt{n-1}}{c_1 \sqrt{\log n}}$$

se obtiene que

$$\frac{1}{\max_i L_{ii}} \sigma \sqrt{(2 + \epsilon) \log n} \leq \frac{c_2 \sigma \sqrt{(2 + \epsilon)}}{\sigma \sqrt{n-1}} < \frac{c_2 \sigma \sqrt{(2 + \epsilon)}}{\sqrt{n-1}} \frac{\sqrt{n-1}}{c_1 \sqrt{\log n}} = \frac{c_2 \sigma \sqrt{2 + \epsilon}}{c_1 \sqrt{\log n}}.$$

De esta manera, reemplazando el término anterior en la desigualdad (2.4.5) obtenemos

$$\begin{aligned} \lambda_{\max}(L) &\leq \left(\sqrt{2 + \epsilon} c_2 + \frac{c_2 \sigma \sqrt{2 + \epsilon}}{c_1 \sqrt{\log n}} \right) \max_i L_{ii} \\ &= \sqrt{2 + \epsilon} c_2 \left(1 + \frac{\sigma}{c_1 \sqrt{\log n}} \right) \max_i L_{ii}. \end{aligned}$$

De la hipótesis (2.4.2) tenemos que para $\epsilon > 0$, $C = c_2 \sqrt{2 + \epsilon}$ cumple

$$\lambda_{\max}(L) \leq C \left(1 + \frac{\sigma}{c_1 \sqrt{\log n}} \right) \max_i L_{ii}$$

con alta probabilidad. □

Cuando el máximo elemento de la diagonal del laplaciano $\max_i L_{ii}$ se comporta como $\sqrt{(n-1) \log n}$ se tiene la condición (2.4.2) además de una cota inferior dada por la siguiente proposición. Este es el caso para la distribución normal y para algunas distribuciones que se pueden aproximar a la misma, como las que se presentan en [6] y que usamos en el ejemplo dos de la Sección 3.

La mayor dificultad es que las entradas de la diagonal del laplaciano no son independientes, pero sin embargo son tales que su matriz de covarianza está dada por

$$\Sigma_L = ((\Sigma_L)_{ij}) = \begin{cases} n-1 & \text{si } i = j \\ 1 & \text{si } i \neq j. \end{cases} \quad (2.4.6)$$

Por lo tanto las correlaciones entre distintos elementos de la diagonal $\rho(L_{ii}, L_{jj})$ tienden a cero cuando $n \rightarrow \infty$. Esta es la razón por la que se puede intuir que asintóticamente el máximo elemento de la diagonal del laplaciano se comporta como el máximo de variables aleatorias independientes.

Observemos que si la matriz Σ_L es de tamaño n , entonces sus valores propios son $2n-2$ con multiplicidad 1 y $n-2$ con multiplicidad $n-1$.

Teorema 2.7. *Sea L una matriz laplaciana cuyas entradas por fuera de la diagonal son independientes con media cero y varianza σ^2 (no dependiendo de n) cumpliendo la Condición 1. Si existe $0 < c < 2$ tal que*

$$\max_i L_{ii} \leq c\sigma\sqrt{(n-1) \log n} \quad (2.4.7)$$

con alta probabilidad, entonces

$$\sqrt{2}\sqrt{\frac{n}{n-1}} \max_i L_{ii} \leq \lambda_{\max}(L) \quad (2.4.8)$$

con alta probabilidad.

Demostración. Consideremos la matriz \hat{L} dada por

$$\hat{L} = \frac{1}{\sigma} L.$$

Notemos que para todo $1 \leq i < j \leq n$ se cumple $\mathbb{E}\hat{L}_{ij} = 0$ y $\mathbb{E}\hat{L}_{ij}^2 = 1$, de esta manera la matriz \hat{L} cumple las hipótesis del Corolario 2.1. Así, para todo $\epsilon > 0$, con

$$\left(2\sqrt{2} - \sqrt{2 + \epsilon}\right) \sqrt{n \log n} \leq \lambda_{\max}(\hat{L})$$

y

$$\sigma \left(2\sqrt{2} - \sqrt{2 + \epsilon}\right) \sqrt{n \log n} \leq \lambda_{\max}(L)$$

con alta probabilidad. Ahora notemos que

$$\begin{aligned} \lambda_{\max}(L) &\geq \sigma \left(2\sqrt{2} - \sqrt{2 + \epsilon}\right) \sqrt{\frac{n}{n-1}} \sqrt{(n-1) \log n} \\ &\geq \sigma \left(2\sqrt{2} - \sqrt{2 + \epsilon}\right) \sqrt{\frac{n}{n-1}} \frac{\max_i L_{ii}}{c\sigma}. \\ &= \frac{1}{c\sqrt{2}} \left(2\sqrt{2} - \sqrt{2 + \epsilon}\right) \sqrt{2} \sqrt{\frac{n}{n-1}} \max_i L_{ii}. \end{aligned}$$

Usando la hipótesis (2.4.7), tomamos $\epsilon > 0$ tal que $2\sqrt{2} - \sqrt{2 + \epsilon} > c\sqrt{2}$, con lo que obtenemos

$$\sqrt{2} \sqrt{\frac{n}{n-1}} \max_i L_{ii} \leq \lambda_{\max}(L) \quad (2.4.9)$$

con alta probabilidad. □

Observación. Las hipótesis (2.4.2) y (2.4.7) se cumplen para el caso de distribución normal y para variables aleatorias acotadas se ha demostrado que (2.4.2) se cumple. A continuación presentamos a detalle el caso de la distribución normal cuya prueba

no fue encontrada en la literatura a pesar de parecer ser un resultado del folklor. El caso de variables acotadas fue recientemente probado en [6]. El caso normal lo presentamos en varios pasos.

Lema 2.2. (a) Sean X_1, X_2, \dots, X_n variables aleatorias simétricas, distintas, no necesariamente con la misma distribución. Entonces

$$\mathbb{P}(X_1 > 0, X_2 > 0, \dots, X_n > 0) > 1 - \frac{1}{2^{n-1}}.$$

(b) En el caso que las variables aleatorias X_1, X_2, \dots, X_n tienen distribución normal multivariada con vector de medias cero y matriz de covarianza definida positiva, se tiene

$$\mathbb{P}(X_1 > 0, X_2 > 0, \dots, X_n > 0) = 1 - \frac{1}{2^n}.$$

Lema 2.3. Sea $L = L_{[X]}$ una matriz $n \times n$ laplaciana con $X = (X_{ij})$ una matriz simétrica con entradas fuera de la diagonal independientes con distribución simétrica. Entonces con alta probabilidad $\max_i L_{ii} > 0$.

Proposición 2.2. Sea $L = L_{[W]}$ una matriz $n \times n$ laplaciana con $W = (W_{ij})$ una matriz simétrica con entradas fuera de la diagonal independientes con distribución normal estándar. Entonces las condiciones (2.4.2) y (2.4.7) se cumplen.

Demostración. Notemos que al ser las entradas de $W = (W_{ij})$ variables aleatorias con distribución normal estándar tenemos

$$\max_i L_{ii} = \sqrt{n-1} \max_i \frac{L_{ii}}{\sqrt{n-1}} = \sqrt{n-1} \max_i \gamma_i,$$

donde γ_i son variables aleatorias con distribución normal estándar.

Ahora, para todo $\delta > 0$, existe $n_0 \in \mathbb{N}$ tal que para $n \geq n_0$,

$$(1 - \delta)\sqrt{\log(n-1)} \leq \max_i \gamma_i \leq (1 + \delta)\sqrt{\log(n-1)},$$

de donde

$$(1 - \delta)\sqrt{(n-1)\log(n-1)} \leq \max_i L_{ii} \leq (1 + \delta)\sqrt{(n-1)\log(n-1)}. \quad (2.4.10)$$

De la desigualdad derecha de (2.4.10) tenemos

$$\max_i L_{ii} \leq (1 + \delta)\sqrt{(n-1)\log(n-1)} \leq (1 + \delta)\sqrt{(n-1)\log(n)},$$

así (2.4.7) se cumple tomando $0 < \delta < 1$.

Por otra parte, de la desigualdad izquierda de (2.4.10),

$$\begin{aligned} \max_i L_{ii} &\geq (1 - \delta)\sqrt{(n-1)\log(n-1)} \\ &= (1 - \delta)\sqrt{\frac{\log(n-1)}{\log(n)}}\sqrt{(n-1)\log(n)} \\ &\geq (1 - \delta)\sqrt{\frac{\log(n_0-1)}{\log(n_0)}}\sqrt{(n-1)\log(n)}. \end{aligned}$$

Así,

$$\sqrt{(n-1)\log(n)} \leq \frac{1}{1 - \delta}\sqrt{\frac{\log(n_0)}{\log(n_0-1)}}\max_i L_{ii}.$$

con lo que se cumple la condición (2.4.2) para $0 < \delta < 1$. \square

Cuando las variables aleatorias $\{W_{ij}\}$ son uniformemente acotadas se cumple la condición (2.4.2) [6].

Observación. Conjeturamos que la Proposición 2.2 también es cierta para variables aleatorias que satisfacen la Condición 1, cuya prueba debiera usar aproximación normal. La mayor dificultad es que los elementos de la diagonal del laplaciano no son independientes, sin embargo debido a la estructura de la matriz de covarianza se pudieran hacer estimaciones adecuadas.

2.5. Sobre la distribución asintótica del mayor eigenvalor

Consideremos $L = L_{[W]}$ una matriz laplaciana cuyas entradas por fuera de la diagonal son variables aleatorias normales estándar independientes.

Sea

$$D^{(n)} = \frac{1}{\sqrt{n-1}}(L_{11}, L_{22}, \dots, L_{nn}),$$

así, $D_i^{(n)}$ tiene distribución normal estándar para $i = 1, \dots, n$. Es importante notar que $\{D_i^{(n)}\}_{i=1, \dots, n}$ no son independientes y por (2.4.6) su matriz de covarianzas está dada por

$$\Sigma_{D^{(n)}} = ((\Sigma_{D^{(n)}})_{ij}) = \begin{cases} 1 & \text{si } i = j \\ \frac{1}{n-1} & \text{si } i \neq j. \end{cases} \quad (2.5.1)$$

Notemos que la matriz $\Sigma_{D^{(n)}}$ tiene por eigenvalores $(n-2)/(n-1)$ con multiplicidad $n-1$ y 2 con multiplicidad 1. Con esto, podemos escribir a $\Sigma_{D^{(n)}}$ como

$$\Sigma_{D^{(n)}} = U\Lambda U^t,$$

donde U es una matriz cuyas columnas son base ortonormal de los vectores propios de Σ_D y Λ es la matriz diagonal con los valores propios de la misma matriz. De esta manera obtenemos que

$$\Sigma_{D^{(n)}}^{1/2} = U\Lambda^{1/2}U^t,$$

donde $\Lambda^{1/2}$ es la matriz diagonal cuyas entradas están dadas por las raíces de los eigenvalores de $\Sigma_{D^{(n)}}$.

Definimos D^* como

$$D^* = \Sigma_{D^{(n)}}^{-1/2} D^{(n)}, \quad (2.5.2)$$

de esta manera, D^* ya tiene por matriz de covarianzas a la matriz identidad y por lo tanto $\{D_i^*\}_{i=1,\dots,n}$ son variables aleatorias normales estándar independientes.

Proposición 2.3. *Sea $L = L_{[W]}$ una matriz $n \times n$ laplaciana con $W = (W_{ij})$ una matriz simétrica con entradas fuera de la diagonal independientes con distribución normal estándar. Sea*

$$a_n = \sqrt{2 \log(n)} \quad y \quad b_n = \sqrt{2 \log(n)} - \frac{\log \log n + \log 4\pi}{2\sqrt{2 \log n}}.$$

Entonces

$$a_n^{-1} \left(\frac{\max_i L_{ii}^{(n)}}{\sqrt{n-1}} - \sqrt{\frac{n-2}{n-1}} b_n \right) \quad (2.5.3)$$

converge en distribución cuando $n \rightarrow \infty$ a una variable aleatoria Gumbel.

Demostración. De la igualdad (2.5.2) tenemos

$$D^{(n)} = \Sigma_{D^{(n)}}^{1/2} D^*,$$

así,

$$D_i^{(n)} = \left(\Sigma_{D^{(n)}}^{1/2} D^* \right)_i = \sum_{j=1}^n \left(\Sigma_{D^{(n)}}^{1/2} \right)_{ij} D_j^*.$$

Ahora,

$$\begin{aligned} \left(\Sigma_{D^{(n)}}^{1/2} \right)_{ij} &= \sum_{l=1}^n \sum_{k=1}^n U_{ik} \Lambda_{kl}^{1/2} U_{lj}^t \\ &= \sum_{l=1}^n U_{il} \Lambda_{ll}^{1/2} U_{lj}^t \\ &= \sum_{l=1}^{n-1} U_{il} \sqrt{\frac{n-2}{n-1}} U_{lj}^t + \sqrt{2} U_{in} U_{nj}^t \\ &= \sqrt{\frac{n-2}{n-1}} \sum_{l=1}^n U_{il} U_{lj}^t + \left(\sqrt{2} - \sqrt{\frac{n-2}{n-1}} \right) U_{in} U_{nj}^t. \end{aligned}$$

Por lo tanto, teniendo en cuenta que las columnas de U son una base ortonormal, obtenemos

$$\left(\Sigma_{D^{(n)}}^{1/2}\right) = \begin{cases} \left(\sqrt{2} - \sqrt{\frac{n-1}{n-2}}\right) U_{in} U_{nj}^t & \text{si } i \neq j \\ \sqrt{\frac{n-2}{n-1}} + \left(\sqrt{2} - \sqrt{\frac{n-1}{n-2}}\right) U_{in} U_{ni}^t & \text{si } i = j. \end{cases} \quad (2.5.4)$$

De esta manera,

$$\begin{aligned} D_i^{(n)} &= \sum_{j=1}^n \left(\Sigma_D^{1/2}\right)_{ij} D_j^* \\ &= \sum_{\substack{j=1 \\ j \neq i}}^n \left(\sqrt{2} - \sqrt{\frac{n-1}{n-2}}\right) U_{in} U_{nj}^t D_j^* + \left(\sqrt{\frac{n-2}{n-1}} + \left(\sqrt{2} - \sqrt{\frac{n-2}{n-1}}\right) U_{in} U_{nj}^t\right) D_i^* \\ &= \left(\sqrt{2} - \sqrt{\frac{n-2}{n-1}}\right) U_{in} \sum_{j=1}^n U_{ni}^t D_j^* + \sqrt{\frac{n-2}{n-1}} D_i^* \end{aligned}$$

Sea

$$Z = Z_n = \sum_{j=1}^n U_n^t D_j^*,$$

como $(D_i^*)_{i=1, \dots, n}$ son variables aleatorias independientes e idénticamente distribuidas normal estándar, Z_n tiene distribución normal con media cero y varianza $\sum_{j=1}^n (U_{nj}^t)^2 = 1$, esto es, Z tiene distribución normal estándar para todo n . Por lo tanto obtenemos que

$$D_i^{(n)} = \left(\sqrt{2} - \sqrt{\frac{n-2}{n-1}}\right) U_{in} Z + \sqrt{\frac{n-2}{n-1}} D_i^*. \quad (2.5.5)$$

Ahora,

$$\begin{aligned}
 a_n^{-1} \left(\max_{1 \leq i \leq n} D_i^{(n)} - \sqrt{\frac{n-2}{n-1}} b_n \right) &= a_n^{-1} \left(\max_{1 \leq i \leq n} D_i^* - b_n \right) \\
 &\quad + a_n^{-1} \left(\max_{1 \leq i \leq n} D_i^{(n)} - \max_{1 \leq i \leq n} D_i^* - \sqrt{\frac{n-2}{n-1}} b_n + b_n \right).
 \end{aligned} \tag{2.5.6}$$

Notemos que el primer término del lado derecho de (2.5.6), converge en distribución a una distribución Gumbel [13]. Por otra parte, usando (2.5.5) obtenemos

$$\begin{aligned}
 \max_{1 \leq i < n} D_i^{(n)} &= \max_{1 \leq i < n} \left(\left(\sqrt{2} - \sqrt{\frac{n-2}{n-1}} \right) U_{in} Z + \sqrt{\frac{n-2}{n-1}} D_i^* \right) \\
 &\leq \left(\sqrt{2} - \sqrt{\frac{n-2}{n-1}} \right) Z \max_{1 \leq i < n} U_{in} + \sqrt{\frac{n-2}{n-1}} \max_{1 \leq i < n} D_i^*.
 \end{aligned}$$

Con lo anterior obtenemos

$$\begin{aligned}
 &\left| \max_{1 \leq i < n} D_i^{(n)} - \max_{1 \leq i < n} D_i^* - \left(\sqrt{\frac{n-2}{n-1}} - 1 \right) b_n \right| \\
 &\leq \left| \left(\sqrt{2} - \sqrt{\frac{n-2}{n-1}} \right) Z \max_{1 \leq i < n} U_{in} \right| + \left| \sqrt{\frac{n-2}{n-1}} \max_{1 \leq i < n} D_i^* - \max_{1 \leq i < n} D_i^* - \left(\sqrt{\frac{n-2}{n-1}} - 1 \right) b_n \right| \\
 &\leq \left(\sqrt{2} - \sqrt{\frac{n-2}{n-1}} \right) \left| Z \max_{1 \leq i < n} U_{in} \right| + \left(\sqrt{\frac{n-2}{n-1}} - 1 \right) \left| \max_{1 \leq i < n} D_i^* - b_n \right|.
 \end{aligned}$$

Luego, como $a_n \rightarrow \infty$ cuando $n \rightarrow \infty$, $\max_{1 \leq i < n} U_{in}$ está acotado puesto que las columnas de matriz U son base ortonormal y por una aplicación del teorema de Slutsky, encontramos que

$$a_n^{-1} \left| \max_{1 \leq i < n} D_i^{(n)} - \max_{1 \leq i < n} D_i^* - \left(\sqrt{\frac{n-2}{n-1}} - 1 \right) b_n \right| \rightarrow 0 \quad \text{cuando} \quad n \rightarrow \infty.$$

Por lo tanto, el segundo término del lado derecho de (2.5.6) tiende a cero cuando $n \rightarrow \infty$. Así, por todo lo anterior encontramos que

$$a_n^{-1} \left(\max_{1 \leq i \leq n} D_i^{(n)} - \sqrt{\frac{n-2}{n-1}} b_n \right)$$

converge en distribución cuando $n \rightarrow \infty$ a una variable aleatoria Gumbel. \square

Observación. El resultado anterior es similar al Teorema 3.1 de [7], sin embargo en éste último se considera la sucesión $\{X_n : n = 0, n = \pm 1, \dots\}$ que es fija, mientras que en el caso nuestro se tiene para cada n la sucesión de arreglos $\{L_{ii}^{(n)} : i = 1, \dots, n\}$.

El siguiente teorema nos dice como es la distribución asintótica del mayor eigenvalor de una matriz laplaciana asociada a una matriz W con entradas normales estándar independientes.

Teorema 2.8. Sea $L = L_{[W]}$ una matriz $n \times n$ laplaciana con $W = (W_{ij})$ una matriz simétrica con entradas fuera de la diagonal independientes con distribución normal estándar. Sea

$$a_n = \sqrt{2 \log(n)} \quad y \quad b_n = \sqrt{2 \log(n)} - \frac{\log \log n + \log 4\pi}{2\sqrt{2 \log n}}.$$

Entonces para todo $C > \sqrt{2}$,

$$\frac{1}{Ca_n} \left(\frac{\lambda_{\max}(L)}{\sqrt{n-1}} - Cb_n \sqrt{\frac{n-2}{n-1}} \right) \quad (2.5.7)$$

converge en distribución cuando $n \rightarrow \infty$ a una variable aleatoria Gumbel.

Demostración. Notemos que

$$\begin{aligned} \frac{1}{Ca_n} \left(\frac{\lambda_{\max}(L)}{\sqrt{n-1}} - Cb_n \sqrt{\frac{n-2}{n-1}} \right) &= \frac{1}{a_n} \left(\max_{1 \leq i \leq n} D_i^{(n)} - b_n \sqrt{\frac{n-2}{n-1}} \right) \\ &\quad + \frac{1}{Ca_n} \left(\frac{\lambda_{\max}(L)}{\sqrt{n-1}} - C \max_{1 \leq i \leq n} D_i^{(n)} \right). \end{aligned} \quad (2.5.8)$$

Por la Proposición 2.3 el primer término de la igualdad derecha de (2.5.8) converge en distribución a una variable aleatoria Gumbel. Por otra parte, del Teorema 2.6 y la Proposición 2.2, tenemos que para $C > \sqrt{2}$

$$\left| \frac{\lambda_{\max}(L)}{\sqrt{n-1}} - C \max_{1 \leq i \leq n} D_i^{(n)} \right| \leq \left| \frac{C}{\sqrt{\log n}} \max_{1 \leq i \leq n} D_i^{(n)} \right|,$$

de esta manera

$$\begin{aligned} \frac{1}{Ca_n} \left| \frac{\lambda_{\max}(L)}{\sqrt{n-1}} - C \max_{1 \leq i \leq n} D_i^{(n)} \right| &\leq \frac{1}{Ca_n} \left| \frac{C}{\sqrt{\log n}} \max_{1 \leq i \leq n} D_i^{(n)} \right| \\ &\leq \frac{1}{\sqrt{\log n}} \frac{1}{a_n} \left| \max_{1 \leq i \leq n} D_i^{(n)} - b_n \sqrt{\frac{n-2}{n-1}} + b_n \sqrt{\frac{n-2}{n-1}} \right| \\ &\leq \frac{1}{\sqrt{\log n}} \frac{1}{a_n} \left| \max_{1 \leq i \leq n} D_i^{(n)} - b_n \sqrt{\frac{n-2}{n-1}} \right| \\ &\quad + \frac{1}{a_n \sqrt{\log n}} b_n \sqrt{\frac{n-2}{n-1}}. \end{aligned} \quad (2.5.9)$$

Luego por el Teorema 2.3 y el Teorema de Slutsky el primer término de la desigualdad anterior tiende a cero conforme $n \rightarrow \infty$. Por otra parte, por como están definidos a_n y b_n tenemos que el segundo término de la desigualdad (2.5.9) tiende a cero cuando $n \rightarrow \infty$. De esta manera

$$\frac{1}{Ca_n} \left| \frac{\lambda_{\max}(L)}{\sqrt{n-1}} - C \max_{1 \leq i \leq n} D_i^{(n)} \right| \rightarrow 0 \quad \text{cuando} \quad n \rightarrow \infty.$$

Por lo tanto, el segundo término del lado derecho de la ecuación (2.5.8) tiende a cero cuando $n \rightarrow \infty$ y por lo tanto obtenemos que para todo $C > \sqrt{2}$,

$$\frac{1}{Ca_n} \left(\frac{\lambda_{\max}(L)}{\sqrt{n-1}} - C b_n \sqrt{\frac{n-2}{n-1}} \right)$$

converge en distribución cuando $n \rightarrow \infty$ a una variable aleatoria Gumbel.

□

Capítulo 3

Aplicaciones

En este capítulo presentamos tres aplicaciones sobre optimización convexa donde el mayor eigenvalor de una matriz laplaciana aleatoria surge de manera natural. Estos ejemplos aparecen en un artículo de Bandeira [6] publicado en 2018. Nuestro enfoque permite extender las tres aplicaciones en diversos sentidos, para lo cual usamos los resultados de optimización convexa del Capítulo 1 y los correspondientes al mayor eigenvalor asintótico de matrices laplacianas aleatorias desarrollado en el Capítulo 2. La Sección 3.1 aborda el llamado problema de \mathbb{Z}_2 sincronización con ruido general. La Sección 3.2 considera el problema de sincronización sobre \mathbb{Z}_2 , también llamada: problema de sincronización corrupta. Finalmente en la Sección 3.3 se estudia el Modelo de bloques estocásticos con 2 comunidades.

3.1. \mathbb{Z}_2 sincronización con ruido general

Dado un nivel de ruido σ y un vector $z \in \{-1, 1\}^n$, supongamos que tenemos las siguientes medidas de ruido

$$Y_{ij} = z_i z_j + \sigma W_{ij}$$

donde $W_{ij} = W_{ji}$ son variables aleatorias independientes con distribución simétrica cumpliendo la Condición 1 del Capítulo 2, con $\mathbb{E}W_{ij} = 0$ y $\mathbb{E}W_{ij}^2 = 1$. El objetivo es recuperar z con alta probabilidad.

Para abordar este problema, usamos el estimador de mínimos cuadrados¹ el cual es la solución del siguiente problema de optimización

$$\begin{aligned} \text{máx} \quad & x^t Y x \\ \text{s.t} \quad & x \in \mathbb{R}^n \\ & x_i^2 = 1. \end{aligned} \tag{3.1.1}$$

Notemos que $x^t Y x = \text{Tr}(x^t Y x) = \text{Tr}(Y x x^t) = \text{Tr}(Y X)$ donde $X = x x^t$, de esta manera el problema (3.1.1) se puede reescribir como

$$\begin{aligned} \text{máx} \quad & \text{Tr}(Y X) \\ \text{s.t} \quad & X_{ii} = 1 \\ & X \succeq 0 \\ & \text{rango}(X) = 1. \end{aligned} \tag{3.1.2}$$

Debido a que el rango no es una función convexa, el problema de optimización anterior no pertenece a la clase de problemas de optimización semidefina. Motivados por [6] usamos el método de *relajamiento* (*relaxation*), que lo que busca es solucionar el problema sin las restricciones no convexas. De esta manera obtenemos

¹En algunos casos, como en el ruido gaussiano, este estimador coincide con el estimador de máxima verosimilitud.

$$\begin{aligned}
& \text{máx} && \text{Tr}(YX) \\
& \text{s.t} && X_{ii} = 1 \\
& && X \succeq 0
\end{aligned} \tag{3.1.3}$$

que si es un problema de programación semidefinida.

Notemos que si $X = xx'$ es la única solución de (3.1.3), entonces x es solución de (3.1.1), con lo que tenemos una manera eficiente de encontrar el estimador de mínimos cuadrados solucionando un problema de programación semidefinida.

El problema dual (3.1.3) está dado por

$$\begin{aligned}
& \text{mín} && \text{Tr}(D) \\
& \text{s.t} && D \text{ diagonal} \\
& && D - Y \succeq 0.
\end{aligned} \tag{3.1.4}$$

Usando el Teorema 1.1 obtenemos que si X y D son soluciones factibles de (3.1.3) y (3.1.4) respectivamente, entonces

$$\text{Tr}(YX) \leq \text{Tr}(D).$$

Por otra parte, al ser X y $D - Y$ semidefinidas positivas, obtenemos

$$0 \leq \text{Tr}((D - Y)X) = \text{Tr}(DX) - \text{Tr}(YX) = \text{Tr}(D) - \text{Tr}(YX).$$

De esta manera si existe una solución factible D para (3.1.4) tal que $\text{Tr}(D) = \text{Tr}(Yxx^t)$, entonces se tiene que $X = xx^t$ es solución óptima de (3.1.3). Por el Lema 1.3, sabemos que para garantizar unicidad de la solución es necesario que $\lambda_2(D - Y) > 0$ y además la solución al problema (3.1.4) esta dada por

$$D = D_{[\text{diag}(x)Y\text{diag}(x)]}.$$

Observemos que por simetría de la distribución de $W = (W_{ij})$, $\text{diag}(z)W\text{diag}(z) \sim W$ (son iguales en distribución) por lo que, sin pérdida de generalidad, podemos tomar $z = \mathbf{1}$. De esta manera tenemos que garantizar que

$$\lambda_2(L_{[Y]}) = \lambda_2(D_{[Y]} - Y) = \lambda_2(D_{[\mathbf{1}\mathbf{1}^t + \sigma W]} - (\mathbf{1}\mathbf{1}^t + \sigma W)) > 0.$$

Notemos que $D_{[\mathbf{1}\mathbf{1}^t + \sigma W]} = nI_n + \sigma D_{[W]}$, por lo que

$$D_{[Y]} - Y = nI_n + \sigma D_{[W]} - \mathbf{1}\mathbf{1}^t - \sigma W = (nI_n - \mathbf{1}\mathbf{1}^t) + \sigma(-D_{[W]} + W) = L_{\mathbf{1}\mathbf{1}^t} + \sigma L_{[W]}.$$

De esta manera, necesitamos que

$$\lambda_2(L_{[Y]}) = \lambda_2(D_{[Y]} - Y) = \lambda_2(L_{\mathbf{1}\mathbf{1}^t} + \sigma L_{[W]}) > 0,$$

pero usando la Proposición 1.2 lo anterior es equivalente a

$$\lambda_{\text{máx}}(L_{[-W]}) < \frac{n}{\sigma}.$$

Por lo tanto, para solucionar el problema de sincronización es necesario determinar cuándo el mayor eigenvalor de la matriz laplaciana asociada $-W$ es menor que n/σ con alta probabilidad, lo cual lo asegura el siguiente resultado.

Proposición 3.1. *Sea $z \in \{-1, +1\}^n$ e $Y = zz^t + \sigma W$ con W una matriz simétrica cuyas entradas fuera de la diagonal son independientes con distribución simétrica con media cero, varianza uno y cumpliendo la Condición 1 del Capítulo 2. Si existe $\epsilon > 0$ tal que*

$$\sigma < \sqrt{\frac{n}{(2 + \epsilon) \log n}},$$

entonces con alta probabilidad $X = zz^t$ es la única solución de (3.1.3).

Demostración. Notemos que W es igual en distribución a $-W$, y cumple las hipótesis del Corolario 2.1. De esta manera para todo $\epsilon > 0$, con alta probabilidad

$$\lambda_{\max}(L_{[-W]}) \leq \sqrt{(2 + \epsilon)n \log(n)}.$$

Tomando $\epsilon > 0$ como el de la hipótesis, obtenemos

$$\sqrt{(2 + \epsilon)n \log n} = n \sqrt{\frac{(2 + \epsilon) \log n}{n}} = \frac{n}{\sqrt{\frac{n}{(2 + \epsilon) \log n}}} < \frac{n}{\sigma}.$$

De esta manera con alta probabilidad $\lambda_{\max}(L_{[-W]}) < n/\sigma$ con lo que garantizamos que la única solución del problema (3.1.3) es $X = zz^t$ con alta probabilidad. \square

Una manera de estimar el mayor eigenvalor de la matriz laplaciana es mediante el mayor elemento de su diagonal. En el caso en que las entradas de W sean variables aleatorias normales independientes tenemos el siguiente resultado, el cual se sigue del Teorema 2.6 y la Proposición 2.2.

Proposición 3.2. *Sea W una matriz simétrica aleatoria cuyas entradas son variables aleatorias normales independientes con $\mathbb{E}W_{ij} = 0$ y $\mathbb{E}W_{ij}^2 = 1$. Entonces con alta probabilidad*

$$\sqrt{2} \sqrt{\frac{n}{n-1}} \max_i (L_{[-W]})_{ii} \leq \lambda_{\max}(L_{[-W]}) \leq 2 \left(1 + \frac{1}{\sqrt{\log n}}\right) \max_i (L_{[-W]})_{ii}.$$

3.2. Sincronización sobre \mathbb{Z}_2

El problema de sincronización sobre \mathbb{Z}_2 es el siguiente: dada una gráfica simple G con n vértices, queremos reconstruir un vector binario $z \in \{+1, -1\}^n$ dado un vector de ruido Y sobre las aristas. De forma más precisa, siguiendo [6], dada una gráfica $G \in \mathcal{G}(n, p)$ del modelo de Erdős-Rényi, tenemos para cada arista $(i, j) \in E(G)$,

$$Y_{ij} = \begin{cases} z_i z_j & \text{con probabilidad } 1 - \epsilon \\ -z_i z_j & \text{con probabilidad } \epsilon, \end{cases} \quad (3.2.1)$$

donde $\epsilon < 1/2$ representa el nivel de ruido y $Y_{ij} = 0$ en caso que (i, j) no sea arista de G .

El objetivo es saber para qué valores de p y ϵ es posible reconstruir z . Para lograr esto, usaremos programación semidefinida para entender cuando $X = z z^t$ es la única solución del problema de optimización (3.2.2) que presentamos más adelante.

Consideremos $H \subset G$ la gráfica tal que la arista (i, j) (que ya es arista de G) es arista de H con probabilidad ϵ . Esta gráfica la llamaremos *gráfica corrupta*, pues tiene aquellas aristas en los cuales no obtenemos sincronización ($Y_{ij} = -z_i z_j$). Se puede escribir a la matriz Y como

$$Y = \text{diag}(z)(A_G - 2A_H)\text{diag}(z).$$

Observación. Notemos que con la representación anterior, para (i, j) arista de G ,

$$Y_{ij} = z_i z_j ((A_G)_{ij} - 2(A_H)_{ij}) = z_i z_j (1 - 2(A_H)_{ij})$$

donde A_G, A_H son las matrices de adyacencia de las gráficas G y H respectivamente. Notemos que con probabilidad $1 - \epsilon$, $(A_H)_{ij} = 0$, de donde $Y_{ij} = z_i z_j$ y con probabilidad ϵ , $(A_H)_{ij} = 1$, de donde $Y_{ij} = -z_i z_j$, que coincide con la definición (3.2.1).

Por todo lo anterior podemos considerar el siguiente problema de programación semidefinida

$$\begin{aligned} \text{máx} \quad & \text{Tr}[\text{diag}(z)(A_G - 2A_H)\text{diag}(z)X] \\ \text{s.t} \quad & X_{ii} = 1 \\ & X \succeq 0. \end{aligned} \quad (3.2.2)$$

Queremos usar el Lema 1.3 tomando $Y = \text{diag}(z)(A_G - 2A_H)\text{diag}(z)$, por lo que necesitamos que el segundo eigenvalor de $D_{[A_G - A_H]} - \text{diag}(z)(A_G - 2A_H)\text{diag}(z)$ sea positivo.

Notemos que

$$\text{diag}(z)(D_{[A_G]} - 2D_{[A_H]})\text{diag}(z) = D_{[A_G]} - 2D_{[A_H]},$$

y por lo tanto

$$\begin{aligned} \lambda_2(D_{[A_G - A_H]} - \text{diag}(z)(A_G - 2A_H)\text{diag}(z)) & \\ &= \lambda_2(D_{[A_G]} - 2D_{[A_H]} - \text{diag}(z)(A_G - 2A_H)\text{diag}(z)) \\ &= \lambda_2(\text{diag}(z)(D_{[A_G]} - 2D_{[A_H]})\text{diag}(z) \\ &\quad - \text{diag}(z)(A_G - 2A_H)\text{diag}(z)) \\ &= \lambda_2(\text{diag}(z)(D_{[A_G]} - 2D_{[A_H]} - A_G + 2A_H)\text{diag}(z)) \\ &= \lambda_2(D_{[A_G]} - 2D_{[A_H]} - A_G + 2A_H) \\ &= \lambda_2(L_G - 2L_H), \end{aligned}$$

donde L_G y L_H son las matrices laplacianas asociadas a las gráficas G y H . Denotemos $L_{Synchron} = L_G - 2L_H$. Observemos que no es tan directa la forma en la cual podamos relacionar el segundo eigenvalor de $L_{Synchron}$ con el mayor eigenvalor de otra matriz laplaciana. Una manera de lograr esta relación es sumando y restando a $L_{Synchron}$ su valor esperado $\mathbb{E}[L_{Synchron}]$. Con esto

$$\mathbb{E}[L_{Synchron}] = \mathbb{E}[D_G - A_G - 2(D_H - A_H)],$$

y dado que $(A_G)_{ij} \sim \text{Ber}(p)$, $(D_G)_{ii} \sim \text{Bin}(n-1, p)$ y condicional a que (i, j) es arista $(A_H)_{ij} \sim \text{Ber}(\epsilon)$, $(D_H)_{ii} \sim \text{Bin}(n-1, \epsilon p)$, tenemos

$$\mathbb{E}(L_{Synch})_{ii} = \mathbb{E}(D_G)_{ii} - 2\mathbb{E}(D_H)_{ii} = (n-1)p - 2\epsilon p(n-1) = (n-1)p(1-2\epsilon),$$

$$\mathbb{E}(L_{Synch})_{ij} = -\mathbb{E}(A_G)_{ij} + 2\mathbb{E}(A_H)_{ij} = -p + 2\epsilon p.$$

Así, realizando manipulaciones algebraicas, obtenemos

$$\mathbb{E}[L_{Synch}] = np(1-2\epsilon)\mathbb{I}_n - p(1-2\epsilon)\mathbf{1}\mathbf{1}^t = p(1-2\epsilon)L_{11^t}.$$

De esta manera,

$$\begin{aligned} 0 < \lambda_2(L_{Synch}) &= \lambda_2(\mathbb{E}[L_{Synch}] - \mathbb{E}[L_{Synch}] + L_{Synch}) \\ &= \lambda_2(p(1-2\epsilon)L_{11^t} - (\mathbb{E}[L_{Synch}] - L_{Synch})), \end{aligned}$$

y por la Proposición 1.2 tenemos que con alta probabilidad

$$\lambda_{\max}(\mathbb{E}[L_{Synch}] - L_{Synch}) < np(1-2\epsilon).$$

3.2.1. Solución del problema

Consideremos $-\mathring{L} = -L_{Synch} + \mathbb{E}L_{Synch}$. El siguiente resultado nos indica cuándo el Problema (3.2.2) tiene como solución única $X = zz^t$.

Teorema 3.1. *Sea $\theta < 1/2$ y $p(1-2\theta)^2 = c$ con $c \in (0, 1)$. Si existe $\epsilon > 0$ tal que $p > \frac{(2+\epsilon)(1-c)\log n}{(1-2\theta)^2 n}$, entonces con alta probabilidad*

$$\lambda_{\max}(-\mathring{L}) \leq n(1-2\theta),$$

y por lo tanto (3.2.2) tiene por solución única $X = zz^t$.

Observación. Si tomamos $c \leq 1/2$, el resultado anterior implica el Teorema 3.9 de Bandeira [6], lo que es interesante dado que nuestra demostración usa una metodología diferente, como se desarrollo en el Capítulo 2.

3.2.2. Demostración del teorema principal

Haremos primero algunas observaciones. Notemos primero que las entradas por fuera de la diagonal de la matriz \mathring{L} están dadas por

$$-\mathring{L}_{ij} = (A_G)_{ij} - 2(A_H)_{ij} - p(1 - 2\theta),$$

así $\mathbb{E} \left[-\mathring{L}_{ij} \right] = 0$ y

$$\begin{aligned} \mathbb{E} \left[-\mathring{L}_{ij}^2 \right] &= \mathbb{E}[(A_G)_{ij} - 2(A_H)_{ij}]^2 - 2p(1 - 2\theta)((A_G)_{ij} - 2(A_H)_{ij}) + p^2(1 - 2\theta)^2 \\ &= \mathbb{E}(A_G)_{ij}^2 - 4\mathbb{E}[(A_G)_{ij}(A_H)_{ij}] + 4\mathbb{E}(A_H)_{ij}^2 \\ &\quad - 2p(1 - 2\theta)(\mathbb{E}(A_G)_{ij} - 2\mathbb{E}(A_H)_{ij}) + p^2(1 - 2\theta)^2 \\ &= p - 4\mathbb{E}[(A_G)_{ij}(A_H)_{ij}] + 4p\theta + 2p(1 - 2\theta)(p - 2p\theta\mathbb{E}(A_H)_{ij}) + p^2(1 - 2\theta)^2 \\ &= p - 4\mathbb{E}[(A_G)_{ij}(A_H)_{ij}] + 4p\theta - p^2(1 - 2\theta)^2. \end{aligned}$$

Observemos también que

$$\begin{aligned} \mathbb{E}[(A_G)_{ij}(A_H)_{ij}] &= \mathbb{E}[(A_G)_{ij}(A_H)_{ij} | (A_G)_{ij} = 1] \mathbb{P}(A_{ij} = 1) \\ &\quad + \mathbb{E}[(A_G)_{ij}(A_H)_{ij} | (A_G)_{ij} = 0] \mathbb{P}(A_{ij} = 0) \\ &= \mathbb{E}[(A_H)_{ij}] \mathbb{P}((A_G)_{ij} = 1) \\ &= \theta p \end{aligned}$$

y de esta manera,

$$\mathbb{E} \left[-\mathring{L}_{ij}^2 \right] = p - p^2(1 - 2\theta)^2.$$

Consideremos ahora $X = (\xi_{ij})$ como la matriz simétrica cuyas entradas están dadas por

$$\xi_{ij} = \frac{-L_{ij}^\circ}{\sqrt{p - p^2(1 - 2\theta)^2}} = \frac{1}{\sqrt{p - p^2(1 - 2\theta)^2}} (-(A_G)_{ij} + 2(A_H)_{ij} + p(1 - 2\theta)).$$

Estas variables aleatorias cumplen $\mathbb{E}\xi_{ij} = 0$, $\text{Var}\xi_{ij} = 1$ y además $\{\xi_{ij}, 1 \leq i < j \leq n\}$ son independientes, pues están dadas por las aristas de la gráfica del modelo de Erdős-Renyi que son independientes. Usando el Corolario 2.1, tenemos que para todo $\epsilon > 0$ con alta probabilidad

$$\lambda_{\max}(L_X) \leq \sqrt{(2 + \epsilon)n \log n}.$$

Observemos que

$$L_X = \frac{1}{\sqrt{p - p^2(1 - 2\theta)^2}} (-\dot{L}),$$

así,

$$\begin{aligned} \lambda_{\max}(-\dot{L}) &\leq \sqrt{(2 + \epsilon)n \log n (p - p^2(1 - 2\theta)^2)} \\ &= n \sqrt{\frac{(2 + \epsilon)(p - p^2(1 - 2\theta)^2) \log n}{n}}. \end{aligned} \quad (3.2.3)$$

De esta forma para asegurar que el problema de sincronización tenga como solución única a $X = zz^t$, tenemos que garantizar que

$$\sqrt{\frac{(2 + \epsilon)(p - p^2(1 - 2\theta)^2) \log n}{n}} \leq p(1 - 2\theta).$$

Demostración (Teorema 3.1). Notemos que $p - p^2(1 - 2\theta)^2 = p(1 - c)$, así

$$\frac{(2 + \epsilon) \log n}{n} \leq \frac{p^2(1 - 2\theta)}{p(1 - c)} = \frac{p^2(1 - 2\theta)}{p - p^2(1 - 2\theta)^2},$$

y de esta manera

$$\frac{(2 + \epsilon) \log n (p - p^2(1 - 2\theta)^2)}{n} \leq p^2(1 - 2\theta)^2.$$

Luego, usando la ecuación (3.2.3), tenemos que con alta probabilidad

$$\lambda_{\max}(-\mathring{L}) \leq n \sqrt{\frac{(2 + \epsilon)(p - p^2(1 - 2\theta)^2) \log n}{n}} \leq n \sqrt{p^2(1 - 2\theta)^2} = np(1 - 2\theta).$$

□

3.2.3. Estimación del mayor eigenvalor

El siguiente teorema para la estimación del mayor eigenvalor se sigue de usar el Teorema 2.6. La condición (2.4.2) se sigue del trabajo de Bandeira (página 375 de [6]). Más específicamente existe una constante $c_2 < 1/\sqrt{2}$ tal que con alta probabilidad

$$\sqrt{(n - 1)(p - p^2(1 - 2\theta)^2) \log n} \leq c_2 \max_i -L_{ii}, \quad (3.2.4)$$

la cual es la condición (2.4.2) con

$$\sigma = (p - p^2(1 - 2\theta)^2). \quad (3.2.5)$$

Además en este caso la constante C del Teorema 2.6 es siempre igual a uno para cualquier $\epsilon > 0$.

Teorema 3.2. *Sea $\theta < 1/2$ y $p(1 - 2\theta)^2 = c$ con $c \in (0, 1)$. Si existe $\epsilon > 0$ tal que $p > \frac{(2+\epsilon)(1-c) \log n}{(1-2\theta)^2 n}$, entonces con alta probabilidad*

$$\lambda_{\max}(-\mathring{L}) \leq \left(1 + \frac{(1 - c)\sqrt{2 + \epsilon}}{(1 + 2\theta)\sqrt{\log(n)}} \right) \max_i -\mathring{L}_{ii}. \quad (3.2.6)$$

Observación. *Pensamos también que se tiene la condición (2.4.7) con σ como en (3.2.5), lo cual, por el Teorema 2.7, nos daría una cota inferior para $\lambda_{\max}(-\mathring{L}_{ii})$.*

3.3. Modelo de bloques estocásticos

Nuestra tercera aplicación también fue estudiada por Bandeira en [6]. Para n par y $0 \leq q < p \leq 1$, decimos que una gráfica aleatoria G es construida del *modelo de bloques estocásticos (MBE)* con dos comunidades, $\mathcal{G}(n, p, q)$, si G tiene n vértices divididos en dos grupos de tamaño $n/2$ y para cada par de vértices i, j , (i, j) es arista de G con probabilidad p si i y j pertenecen al mismo grupo y con probabilidad q en otro caso independiente de cualquier otra arista. Consideramos que p y q son de la forma

$$p = p_n := \frac{\alpha \log n}{n}, \quad q = q_n := \frac{\beta \log n}{n} \quad \text{con } \alpha > \beta > 0. \quad (3.3.1)$$

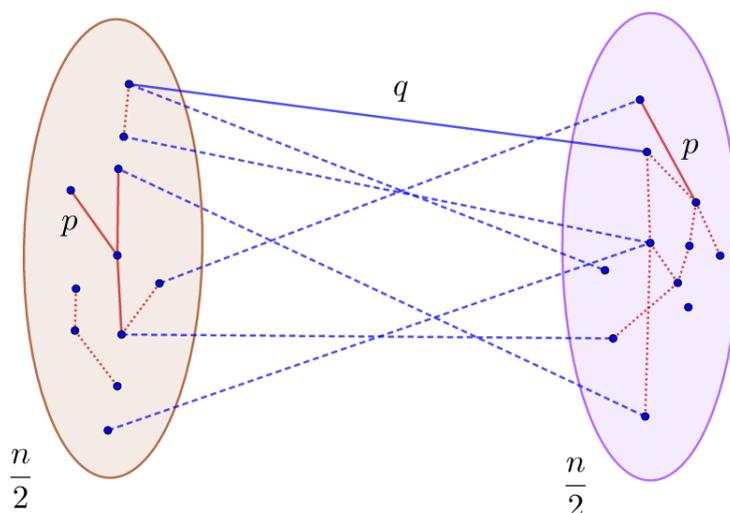


Figura 3.1: Modelo de bloques estocásticos con dos comunidades

Sea $z \in \{-1, 1\}^n$ el vector tal que es igual a 1 en uno de los grupos y -1 en el otro. Esto es, z es el vector que nos da información sobre qué vértice pertenece a cada grupo. El objetivo es recuperar z a partir de G .

Sea B la matriz de adyacencia con signo de G , esto es, $B_{ij} = 1$ si (i, j) es arista de G , $B_{ij} = -1$ si no lo es y $B_{ii} = 0$. Notemos que para cada vector $x \in \{-1, +1\}^n$, $(x^t B x)_{ij} = x_i x_j B_{ij}$. De esta manera si $B_{ij} = 1$ entonces (i, j) es arista y queremos determinar si esa arista está en un grupo. Así, su valor será máximo si $x_i = x_j$ es decir, los vértices i, j pertenecen al mismo grupo. Por otra parte, si (i, j) no es arista entonces $B_{ij} = -1$, de manera que el valor máximo de $x_i x_j B_{ij}$ será cuando $x_i \neq x_j$, es decir cuando los vértices i, j están en distintos grupo. Por lo tanto, $\max x^t B x$ será la configuración que maximice la probabilidad de obtener el vector z correcto.

Observación. Otra punto de vista es que tenemos dos tipos de aristas (con vértices en un mismo grupo y con vértices en grupos diferentes), la primera con probabilidad p y la segunda con probabilidad $q < p$. Por tanto es más probable que las aristas estén formadas por vértices que pertenezcan a un mismo grupo por lo que $x_i x_j B_{ij}$ es máximo en estas condiciones. Por otro lado, dos vértices en el mismo grupo no están unidos con probabilidad $1 - p$ y si están en distintos grupos no están unidos con probabilidad $1 - q > 1 - p$. Así es más probable que para dos vértices que no forman una arista, éstos están en grupos diferentes.

Se tiene entonces el siguiente problema de optimización

$$\begin{aligned}
 \text{máx} \quad & x^t B x \\
 \text{s.t} \quad & x \in \mathbb{R}^n \\
 & x_1^2 = 0 \\
 & \sum_{i=1}^n x_i = 0.
 \end{aligned} \tag{3.3.2}$$

Notemos que la matriz B la podemos escribir como $B = 2A - (\mathbf{1}\mathbf{1}^t - I)$ donde A es la matriz de adyacencia de la gráfica G . Siguiendo el mismo enfoque que en la sección anterior, usaremos el método de relajamiento, con lo que obtenemos el siguiente problema de programación semidefinida

$$\begin{aligned}
& \text{máx} \quad \text{Tr}((2A - (\mathbf{1}\mathbf{1}^t - I))X) \\
& \text{s.t} \quad X_{ii} = 1 \\
& \quad \quad X \succeq 0.
\end{aligned} \tag{3.3.3}$$

Usando el Lema 1.3 obtenemos que $X = zz^t$ es la única solución al problema (3.3.3) siempre que el segundo eigenvalor de

$$D_{[\text{diagz}(2A - (\mathbf{1}\mathbf{1}^t - I))\text{diagz}]} - (2A - (\mathbf{1}\mathbf{1}^t - I))$$

sea estrictamente positivo.

Para resolver esto definimos la siguiente matriz

$$\Gamma_{MBE} = D_+ - D_- - A,$$

donde D_+ es la matriz diagonal de grados internos, es decir, es el número de vecinos que un vértice tiene dentro de un mismo grupo, D_- es la matriz diagonal de grados externos, es decir, es el número de vecinos que un vértice tiene con vértices del otro grupo y A es la matriz de adyacencia de G . Es importante observar que Γ_{MBE} no es una matriz laplaciana.

Sin pérdida de generalidad, etiquetaremos los vértices de tal manera que los primeros $n/2$ pertenezcan a un mismo grupo.

Notemos que

$$D_{[\text{diagz}(2A - (\mathbf{1}\mathbf{1}^t - I))\text{diagz}]} - (2A - (\mathbf{1}\mathbf{1}^t - I)) = 2\Gamma_{MBE} + \mathbf{1}\mathbf{1}^t.$$

De esta manera, para garantizar que zz^t sea la única solución al problema (3.3.3) es necesario que $\lambda_2(2\Gamma_{MBE} + \mathbf{1}\mathbf{1}^t) > 0$.

Como en el caso de la sección anterior, no es tan fácil relacionar el segundo eigenvalor de $2\Gamma_{MBE} - \mathbf{1}\mathbf{1}^t$ con el mayor eigenvalor de cierta matriz, por lo que usaremos el mismo argumento de sumar y restar el valor esperado.

Recordemos que $(D_+)_{ii} \sim \text{Bin}(n/2 - 1, p)$, $(D_-)_{ii} \sim \text{Bin}(n/2, q)$ y $A_{ij} \sim \text{Ber}(p)$ si i, j pertenecen a un mismo grupo y $A_{ij} \sim \text{Ber}(q)$ en caso contrario. Así,

$$\mathbb{E}(\Gamma_{MBE})_{ii} = \mathbb{E}((D_+)_{ii} - (D_-)_{ii}) = \left(\frac{n}{2} - 1\right)p - \frac{n}{2}q$$

y

$$\mathbb{E}(\Gamma_{MBW})_{ij} = -\mathbb{E}A_{ij} = \begin{cases} -p & \text{si } i, j \text{ pertenecen a un mismo grupo} \\ -q & \text{en caso contrario.} \end{cases}$$

De esta manera, luego de manipulaciones algebraicas y dado que el vector z tiene información sobre qué vértices pertenecen a cada grupo, se obtiene que

$$\mathbb{E}\Gamma_{MBE} = \left(\frac{n}{2}p - \frac{n}{2}q\right) \mathbb{I}_n - \left(\frac{p+q}{2} \mathbf{1}\mathbf{1}^t + \frac{p-q}{2} z z^t\right).$$

Para garantizar que $z z^t$ sea la única solución del problema (3.3.2) es necesario que $\lambda_{\max}(-\Gamma_{MBE} + \mathbb{E}\Gamma_{MBE}) < \frac{n}{2}(p - q)$ lo cual es el Lema 3.12 de [6].

3.3.1. Solución del problema

Sean α y β como en (3.3.1). El siguiente resultado da la solución al problema (3.3.2). Este resultado lo estableció Bandeira en [6], pero nuestra prueba es diferente en dos sentidos. El primero es que usamos nuestro Corolario 2.1 consecuencia del trabajo de Ding-Jiang [9], así mismo nuestro enfoque sugiere cómo se podría generalizar el resultado a más de dos comunidades. Igualmente pensamos que nuestras pruebas son más sencillas usando resultados conocidos sobre el mayor eigenvalor de matrices laplacianas y de adyacencia, y otros como resultados de desviaciones grandes.

Teorema 3.3. *Sea $n \geq 4$ par y sea $G \in \mathcal{G}(n, p, q)$ con probabilidades*

$$p = p_n := \frac{\alpha \log n}{n}, \quad q = q_n := \frac{\beta \log n}{n} \quad \text{con } \alpha > \beta > 0.$$

Si

$$\sqrt{2} < \sqrt{\alpha} - \sqrt{\beta} \quad (3.3.4)$$

y además existe $0 < \epsilon < 1$ tal que

$$\sqrt{\alpha}(\sqrt{2(1+\epsilon)} - 1) < \sqrt{\beta}, \quad (3.3.5)$$

entonces con alta probabilidad

$$\lambda_{\max}(-\Gamma_{MBE} + \mathbb{E}\Gamma_{MBE}) < \frac{n}{2}(p - q),$$

y por lo tanto (3.3.2) tiene solución única $X = zz^t$ con alta probabilidad.

Cabe resaltar que en el trabajo de Bandeira [6] no se requiere la condición (3.3.5), pero en nuestra prueba es esencial.

3.3.2. Demostración del teorema principal

Necesitamos encontrar cotas para el mayor eigenvalor de la matriz Γ_{MBE} con el fin de determinar como deben ser las probabilidades p y q para que el problema de sincronización (3.3.2) tenga solución única con alta probabilidad. Una dificultad es que en este caso no tenemos la misma varianza para todas las entradas por fuera de la diagonal de Γ_{MBE} . Por esta razón, descompondremos esta matriz como diferencia de dos matrices laplacianas en las cuales podremos usar los resultados de la Sección 2.

Recordemos que $\Gamma_{MBE} = D_+ - D_- - A$ no es una matriz laplaciana. Para solucionar esto, sea g el vector que es 1 en uno de los grupos y -1 en el otro, de esta manera, la matriz $\Gamma'_{MBE} = \text{diag}(g)\Gamma_{MBE}\text{diag}(g)$ si es una matriz laplaciana.

Notemos que

$$\Gamma'_{MBE} = D_+ - D_- - \text{diag}(g)A\text{diag}(g),$$

por lo que,

$$(\Gamma'_{MBE})_{ij} = -g_i g_j A_{ij}, \quad (\Gamma'_{MBE})_{ii} = (D_+)_{ii} - (D_-)_{ii}.$$

Como g nos da información sobre el grupo, si $g_i g_j = 1$ tenemos que los vértices i, j pertenecen al mismo grupo, mientras que si $g_i g_j = -1$ entonces los vértices están en distintos grupos. Con esta observación tenemos que la matriz $\text{diag}(g)A\text{diag}(g)$ se puede descomponer como la suma de dos matrices $-A_+, A_-$, donde A_+ es la matriz de adyacencia asociada a los vértices que están en el mismo grupo y A_- la matriz de adyacencia asociada a los vértices en grupo distintos. De esta manera obtenemos que

$$\Gamma'_{MBE} = D_+ - A_+ - D_- + A_- = L_+ - L_-$$

es decir, Γ'_{MBE} es la diferencia de dos matrices laplacianas donde L_+ corresponde a una gráfica con dos componentes de tamaño $n/2$ mientras que L_- corresponde a una gráfica bipartita².

La matriz L_+ es una matriz a dos bloques, los cuales están asociadas a gráficas de Erdős-Rényi de tamaño $n/2$. Consideremos las entradas por fuera de la diagonal de L_+ como $\xi_{ij} \sim \text{Ber}(p)$ y para L_- como $\eta_{ij} \sim \text{Ber}(q)$ y sea $\mathring{L} = L - \mathbb{E}L$.

Sea

$$\begin{aligned} -\Gamma'_{MBE} \circ &= -\Gamma'_{MBE} + \mathbb{E}\Gamma'_{MBE} \\ &= -L_+ + L_- + \mathbb{E}[L_+ - L_-] \\ &= -\mathring{L}_+ + \mathring{L}_-. \end{aligned}$$

Notemos que en el evento $C_{ij} = \{\text{los vértices } i, j \text{ pertenecen al mismo grupo}\}$ obtenemos

$$(-\Gamma'_{MBE})_{ij} \circ = (-\mathring{L}_+)_{ij} \mathbb{I}_{C_{ij}} + (\mathring{L}_-)_{ij} \mathbb{I}_{C_{ij}^c} = (-\xi_{ij} + p) \mathbb{I}_{C_{ij}} + (\eta_{ij} - q) \mathbb{I}_{C_{ij}^c}.$$

²Una gráfica bipartita es aquella cuyos vértices se pueden separar en dos grupos disjuntos de forma que sólo se tengan aristas entre estos grupos (y no dentro del mismo).

Ahora, usando desigualdad triángular (Lema 1.1) obtenemos que

$$\lambda_{\max}(-\Gamma'_{MBE}) = \lambda_{\max}(-\overset{\circ}{L}_+ + \overset{\circ}{L}_-) \leq \lambda_{\max}(-\overset{\circ}{L}_+) + \lambda_{\max}(\overset{\circ}{L}_-). \quad (3.3.6)$$

En las siguientes dos subsecciones, presentamos estimaciones para los dos términos del lado derecho de (3.3.6). Para el primero de ellos usamos nuestro Teorema 2.5. Para el segundo técnicas de desviaciones grandes así como del comportamiento asintótico del mayor eigenvalor de una matriz de adyacencia (Lema 2.1).

Estimación de $\lambda_{\max}(-\overset{\circ}{L}_+)$

Notemos que $-\overset{\circ}{L}_+ = -L_+ + \mathbb{E}L_+$ solo depende de los vértices que están en cada grupo (sin contar interacciones entre ellos). Podemos entonces usar nuestro Teorema 2.5 con $k = 2$ bloques, obteniendo que para todo $\epsilon > 0$ con alta probabilidad

$$\lambda_{\max}(-\overset{\circ}{L}_+) \leq \sqrt{(1 + \epsilon)p(1 - p)n \log n}. \quad (3.3.7)$$

Estimación de $\lambda_{\max}(\overset{\circ}{L}_-)$

Notemos que en este caso la matriz sólo depende de las aristas entre grupos, por lo que se puede entender como matrices laplacianas de gráficas bipartitas. Observemos que

$$\overset{\circ}{L}_- = L_- - \mathbb{E}L_- = D_- - A_- - \mathbb{E}D_- + \mathbb{E}A_- = D_- - \mathbb{E}D_- + \mathbb{E}A_- - A_-$$

y que $D_- - \mathbb{E}D_-$ es una matriz diagonal cuyas entradas son variables aleatorias $\text{Bin}(n/2, q)$ centradas. Por otra parte $\mathbb{E}A_- - A_-$ es una matriz a bloques cuyas entradas no nulas están por fuera de la diagonal.

De todo lo anterior, usando desigualdad triángular (Lema 1.1) obtenemos que

$$\lambda_{\max}(L_-) \leq \lambda_{\max}(D_- - \mathbb{E}D_-) + \lambda_{\max}(\mathbb{E}A_- - A_-). \quad (3.3.8)$$

Procedamos ahora a estimar los términos del lado derecho de (3.3.8), lo cual hacemos de la siguiente manera

Estimación de $\lambda_{\max}(D_- - \mathbb{E}D_-)$

Notemos que al ser $D_- - \mathbb{E}D_-$ una matriz diagonal, el mayor eigenvalor será el máximo elemento de la diagonal, esto es

$$\lambda_{\max}(D_- - \mathbb{E}D_-) = \max_{1 \leq i \leq n} ((D_- - \mathbb{E}D_-)_{ii}) = \max_{1 \leq i < n} \left((D_-)_{ii} - \frac{n}{2}q \right).$$

Es importante notar que $\{(D_-)_{ii} - nq/2\}_{1 \leq i \leq n}$ no son independientes entre si, pero las primeras $n/2$ de ellas si los son, ésto debido a la estructura de la matriz L_- . Si bien el siguiente resultado tiene una demostración sencilla, no lo encontramos en la literatura.

Proposición 3.3. *Sea $Z = \max_{1 \leq i \leq n/2} \{X_{ii} - nq/2\}$ donde $\{X_{ii}\}_{i=1, \dots, n/2}$ son variables aleatorias independientes e idénticamente distribuidas $\text{Bin}(n/2, q)$, entonces para todo $t > 0$,*

$$\mathbb{P}(Z > t) \leq \exp f(n, t),$$

donde

$$f(n, t) = - \left(t + \frac{n}{2}q \right) \log \left(\frac{4 \left(t + \frac{n}{2}q \right) (1 - q)}{q \left(n^2 - 4 \left(t + \frac{n}{2}q \right) \right)} \right) + \frac{n^2}{4} \log \left(\frac{n^2(1 - q)}{n^2 - 4 \left(t + \frac{n}{2}q \right)} \right).$$

Demostración. Notemos que

$$Z = \max_{1 \leq i \leq n/2} \{X_{ii} - nq/2\} = \max_{1 \leq i \leq n/2} \{X_{ii}\} - nq/2 = \hat{Z} - nq/2,$$

de esta manera obtenemos que para todo $t > 0$ y $\rho > 0$

$$\mathbb{P}(Z > t) = \mathbb{P}\left(\hat{Z} > t + \frac{nq}{2}\right) = \mathbb{P}\left(\rho\hat{Z} > \rho\left(t + \frac{nq}{2}\right)\right) = \mathbb{P}\left(e^{\rho\hat{Z}} > e^{\rho y}\right),$$

donde $y = t + nq/2$. Ahora, usando desigualdad de Markov, independencia y la función generadora de momentos de una variable aleatoria binomial, obtenemos

$$\begin{aligned} \mathbb{P}\left(e^{\rho\hat{Z}} > e^{\rho y}\right) &\leq e^{-\rho y} \mathbb{E} e^{\rho\hat{Z}} \\ &= e^{-\rho y} \mathbb{E} \exp\left(\rho \max_{1 \leq i \leq n/2} X_{ii}\right) \\ &\leq e^{-\rho y} \mathbb{E} \exp\left(\rho \sum_{i=1}^{n/2} X_{ii}\right) \\ &= e^{-\rho y} \prod_{i=1}^{n/2} \mathbb{E} \exp(\rho X_{ii}) \\ &= e^{-\rho y} \prod_{i=1}^{n/2} (e^{\rho q} + 1 - q)^{n/2} \\ &= e^{-\rho y} (e^{\rho q} + 1 - q)^{n^2/4} \\ &= \exp\left(-\rho y + \frac{n^2}{4} \log(e^{\rho q} + 1 - q)\right). \end{aligned}$$

Sea $g(\rho) = -\rho y + \frac{n^2}{4} \log(e^{\rho q} + 1 - q)$, notemos que

$$g'(\rho) = -y + \frac{n^2 e^{\rho q}}{4(e^{\rho q} + 1 - q)},$$

luego, igualando a cero obtenemos

$$\begin{aligned}\frac{n^2 e^\rho q}{4(e^\rho q + 1 - q)} &= y \\ n^2 e^\rho q &= 4y(e^\rho q + 1 - q) \\ e^\rho q(n^2 - 4y) &= 4y(1 - q) \\ e^\rho &= \frac{4y(1 - q)}{q(n^2 - 4y)}.\end{aligned}$$

Por lo tanto obtenemos que el punto crítico de la función $g(\rho)$ es

$$\rho_c = \log \left(\frac{4y(1 - q)}{q(n^2 - 4y)} \right). \quad (3.3.9)$$

Ahora, notemos que la segunda derivada de $g(\rho)$ está dada por

$$g''(\rho) = \frac{n^2 (e^\rho q (e^\rho q + 1 - q) - e^\rho q e^\rho q)}{4(e^\rho q + 1 - q)^2} = \frac{n^2(1 - q)}{4(e^\rho + 1 - q)^2} > 0,$$

por lo tanto (3.3.9) es un mínimo, de donde

$$\begin{aligned}g(\rho_c) &= -y \log \left(\frac{4y(1 - q)}{q(n^2 - 4y)} \right) + \frac{n^2}{4} \log \left(q \frac{4y(1 - q)}{q(n^2 - 4y)} + 1 - q \right) \\ &= -y \log \left(\frac{4y(1 - q)}{q(n^2 - 4y)} \right) + \frac{n^2}{4} \log \left(\frac{4y(1 - q)}{n^2 - 4y} + 1 - q \right) \\ &= -y \log \left(\frac{4y(1 - q)}{q(n^2 - 4y)} \right) + \frac{n^2}{4} \log \left(\frac{(1 - q)(4y + n^2 - 4y)}{n^2 - 4y} \right) \\ &= -y \log \left(\frac{4y(1 - q)}{q(n^2 - 4y)} \right) + \frac{n^2}{4} \log \left(\frac{n^2(1 - q)}{n^2 - 4y} \right).\end{aligned}$$

Así obtenemos

$$\mathbb{P}(Z > t) \leq \exp \left(-y \log \left(\frac{4y(1 - q)}{q(n^2 - 4y)} \right) + \frac{n^2}{4} \log \left(\frac{n^2(1 - q)}{n^2 - 4y} \right) \right).$$

Reemplazando $y = t + nq/2$ se sigue el resultado. \square

Corolario 3.1. Sea $f(n, t)$ como en la Proposición 3.3. Si t es del orden $\log(n)/n$, entonces $f(n, t) < 0$ cuando $t \rightarrow \infty$.

Corolario 3.2. Sea $Z = \max_{1 \leq i \leq n/2} \{X_{ii} - nq/2\}$ donde $\{X_{ii}\}_{i=1, \dots, n/2}$ son variables aleatorias independientes e idénticamente distribuidas $Bin(n/2, q)$. Si t es del orden $\log(n)/n$, entonces con alta probabilidad

$$\lambda_{\max}(D^- - \mathbb{E}D^-) \leq \frac{\log n}{n}. \quad (3.3.10)$$

Estimación de $\lambda_{\max}(\mathbb{E}A_- - A_-)$

Notemos que $\mathbb{E}A_- - A_-$ es de la forma

$$\mathbb{E}A_- - A_- = \left[\begin{array}{c|c} 0 & E \\ \hline E^* & 0 \end{array} \right]. \quad (3.3.11)$$

Además,

$$\left[\begin{array}{c|c} 0 & E \\ \hline E^* & 0 \end{array} \right] \left[\begin{array}{c|c} 0 & E \\ \hline E^* & 0 \end{array} \right] = \left[\begin{array}{c|c} EE^* & 0 \\ \hline 0 & E^*E \end{array} \right]. \quad (3.3.12)$$

De esta manera, los eigenvalores de $\mathbb{E}A_- - A_-$ serán las raíces cuadradas de los eigenvalores de EE^* . Notemos que E es una matriz cuadrada de tamaño $n/2$ con

$$E_{ij} = q - \eta_{ij},$$

donde $\eta_{ij} \sim Ber(q)$, así $\mathbb{E}E_{ij} = 0$ y $\mathbb{E}E_{ij}^2 = q(1 - q)$. Consideremos la matriz $U = (u_{ij})$ dada por

$$U_{ij} = \frac{1 - \eta_{ij}}{\sqrt{q(1 - q)}},$$

así, $\mathbb{E}u_{ij} = 0$, $\mathbb{E}u_{ij}^2 = 1$ y $E = \sqrt{q(1-q)}U$. Es importante notar que la distribución empírica espectral de $2UU^*/n$ converge en distribución a la distribución de Marčenko-Pastur con soporte en $[0, 4]$ (Teorema 2.5 [5]) además el mayor eigenvalor de esta matriz converge casi seguramente a 4 (Teorema 2 de [3]). Luego, para todo $\epsilon > 0$

$$\lambda_{\max}(U) \leq (2 + \epsilon)\sqrt{\frac{n}{2}}.$$

y por lo tanto

$$\lambda_{\max}(\mathbb{E}A_- - A_-) = \lambda_{\max}(E) \leq (2 + \epsilon)\sqrt{q(1-q)\frac{n}{2}}. \quad (3.3.13)$$

Estimación de la cota superior de $\lambda_{\max}(-\Gamma'_{SMB})$

Con lo anterior tenemos los elementos para concluir la demostración del Teorema 3.3. De las desigualdades (3.3.7)-(3.3.13) obtenemos que

$$\begin{aligned} \lambda_{\max}(-\Gamma'_{MBS}) &\leq \sqrt{(1+\epsilon)p(1-p)n \log n} + \frac{\log n}{n} + (2+\epsilon)\sqrt{q(1-q)\frac{n}{2}} \\ &= \sqrt{(1+\epsilon)p(1-p)n \log n} + \frac{\log n}{n} + \frac{1}{\sqrt{2}}(2+\epsilon)\sqrt{q(1-q)n} \\ &= \frac{n}{2} \left(2\sqrt{(1+\epsilon)p(1-p)\frac{1}{n} \log n} + \frac{2 \log n}{n^2} + \frac{2}{\sqrt{2}}(2+\epsilon)\sqrt{\frac{q(1-q)}{n}} \right) \\ &= \frac{n}{2} \left(2\sqrt{\frac{(1+\epsilon)p(1-p)}{n} \log n} + \frac{2 \log n}{n^2} + \sqrt{2}(2+\epsilon)\sqrt{\frac{q(1-q)}{n}} \right). \end{aligned}$$

Recordando que $p = \frac{\alpha \log n}{n}$ y $q = \frac{\beta \log n}{n}$, obtenemos

$$\begin{aligned}\sqrt{\frac{(1+\epsilon)p(1-p)}{n} \log n} &= \sqrt{\frac{(1+\epsilon)}{n} \frac{\alpha \log n}{n} \left(1 - \frac{\alpha \log n}{n}\right) \log n} \\ &= \frac{\log n}{n} \sqrt{\alpha(1+\epsilon) \left(1 - \frac{\alpha \log n}{n}\right)},\end{aligned}$$

y

$$\begin{aligned}\sqrt{2}(2+\epsilon) \sqrt{\frac{q(1-q)}{n}} &= \sqrt{2}(2+\epsilon) \sqrt{\frac{1}{n} \frac{\beta \log n}{n} \left(1 - \frac{\beta \log n}{n}\right)} \\ &= \sqrt{2}(2+\epsilon) \frac{\log n}{n} \sqrt{\frac{\beta}{\log n} \left(1 - \frac{\beta \log n}{n}\right)}.\end{aligned}$$

Así

$$\begin{aligned}\lambda_{\max}(\Gamma'_{MBE}) &\leq \frac{n \log n}{2} \frac{1}{n} \left[2 \sqrt{\alpha(1+\epsilon) \left(1 - \frac{\alpha \log n}{n}\right)} + \frac{2}{n} \right. \\ &\quad \left. + \sqrt{2}(2+\epsilon) \sqrt{\frac{\beta}{\log n} \left(1 - \frac{\beta \log n}{n}\right)} \right].\end{aligned}$$

Como

$$2 \sqrt{\alpha(1+\epsilon) \left(1 - \frac{\alpha \log n}{n}\right)} \leq 2 \sqrt{\alpha(1+\epsilon)}, \quad (3.3.14)$$

entonces, si garantizamos que con alta probabilidad

$$2 \sqrt{\alpha(1+\epsilon)} + \frac{2}{n} + \sqrt{2}(2+\epsilon) \sqrt{\frac{\beta}{\log n} \left(1 - \frac{\beta \log n}{n}\right)} \leq \alpha - \beta, \quad (3.3.15)$$

obtenemos que

$$\lambda_{\max}(\Gamma'_{MBE}) \leq \frac{n \log n}{2} \frac{1}{n} (\alpha - \beta) = \frac{n}{2} (p - q)$$

y por lo tanto, el problema (3.3.2) tiene solución única.

Demostración (Teorema 3.3). De la estimación de $\lambda_{\max}(-\mathring{L}_+)$, más precisamente las ecuaciones (3.3.7) y (3.3.14), tenemos que para todo $\epsilon > 0$, existe $n'_0 \in \mathbb{N}$ tal que

$$\mathbb{P} \left(\lambda_{\max}(-\mathring{L}_1) < \frac{n \log n}{2n} \sqrt{2\alpha(2+\epsilon)} \right) \geq 1 - \epsilon. \quad (3.3.16)$$

Por hipótesis tenemos que existe $0 < \epsilon < 1$ tal que

$$\sqrt{\alpha}(\sqrt{2(1+\epsilon)} - 1) < \sqrt{\beta},$$

de esta manera

$$\frac{\sqrt{2\alpha(1+\epsilon)}}{\sqrt{\alpha} + \sqrt{\beta}} = \kappa < 1$$

cumpliendo (3.3.16) para algún $n'_0 \in \mathbb{N}$.

Así, usando la hipótesis $\sqrt{2} < \sqrt{\alpha} - \sqrt{\beta}$

$$2\sqrt{\alpha(1+\epsilon)} = \sqrt{2}\sqrt{2\alpha(1+\epsilon)} < (\sqrt{\alpha} - \sqrt{\beta})\kappa(\sqrt{\alpha} + \sqrt{\beta}) = \kappa(\alpha - \beta) \quad (3.3.17)$$

Ahora, sea $n''_0 \in \mathbb{N}$ tal que para $n > n''_0$,

$$\frac{2}{n} + \left(\sqrt{2} + \frac{1}{2} \right) (2+\epsilon) \sqrt{\frac{\beta}{\log n} \left(1 - \frac{\beta \log n}{n} \right)} \leq (1 - \kappa)(\alpha - \beta). \quad (3.3.18)$$

De esta manera, para $n_0 > \max\{n'_0, n''_0\}$, de las ecuaciones (3.3.17) y (3.3.18) se sigue (3.3.15). □

Para finalizar realizaremos tres observaciones sobre el rango de valores factibles de α y β .

Observación. Por hipótesis sabemos que $\sqrt{\alpha} - \sqrt{\beta} > \sqrt{2}$, entonces $\sqrt{\beta} < \sqrt{\alpha} - \sqrt{2}$, así reemplazando en la desigualdad (3.3.5)

$$\begin{aligned}\sqrt{\alpha}(\sqrt{2(1+\epsilon)} - 1) &< \sqrt{\alpha} - \sqrt{2} \\ \sqrt{2} &< \sqrt{\alpha}(2 - \sqrt{2(1+\epsilon)}) \\ \frac{\sqrt{2}}{2 - \sqrt{2(1+\epsilon)}} &< \sqrt{\alpha}.\end{aligned}\tag{3.3.19}$$

Como $\epsilon > 0$, se tiene la desigualdad $2 - \sqrt{2(1+\epsilon)} < 2 - \sqrt{2}$, así obtenemos

$$1 + \sqrt{2} = \frac{\sqrt{2}}{2 - \sqrt{2}} < \sqrt{\alpha},\tag{3.3.20}$$

que nos da una cota inferior para α .

Observación. De la desigualdad (3.3.5) tenemos que

$$\alpha < \frac{\beta}{(\sqrt{2(1+\epsilon)} - 1)^2} < \frac{\beta}{(\sqrt{2} - 1)^2} = (3 + 2\sqrt{2})\beta,\tag{3.3.21}$$

de donde encontramos una cota superior para α . Recordemos que α y β nos determinan como son las probabilidades p_n, q_n , así con la desigualdad anterior obtenemos

$$q_n < p_n < (3 + 2\sqrt{2})q_n.$$

Observación. De la desigualdad (3.3.21) tenemos

$$\alpha - \beta < \alpha_{\text{máx}} - \beta = (2 + 2\sqrt{2})\beta.$$

Por otra parte, de la cota (3.3.20) obtenemos

$$1 + \sqrt{2} - \beta < \alpha - \beta,$$

como por hipótesis $\alpha - \beta > 0$, de todo lo anterior encontramos que

$$\text{máx}\{1 + \sqrt{2} - \beta, 0\} < \alpha - \beta < (2 + \sqrt{2})\beta.$$

Apéndice A

Reporte de simulaciones

En este apéndice presentamos resultados de un estudio de simulación sobre la distribución empírica espectral de matrices laplacianas con distintas distribuciones y tamaños y sobre la distribución del mayor eigenvalor de matrices laplacianas con distribución normal. Estos resultados nos llevaron a plantear las conjeturas que presentamos en el Capítulo 2 sobre una aproximación a la densidad de la convolución libre $\gamma_M = S \boxplus \gamma$ de la distribución semicircular S y la distribución normal estándar γ mediante una mezcla de las distribuciones S y γ .

Para el primer caso, simulamos matrices $X = (X_{ij})$ de cuatro diferentes tamaños ($n = 100, 500, 1000, 2000$) y tres distribuciones simétricas (normal, Bernoulli, uniforme) con media cero y varianza uno. Definimos la matriz laplaciana $L = L_{[X]}$ para cada matriz X . Elegimos estas distribuciones para tener ejemplos de variables continuas y discretas y con soporte acotado y no acotado. En todos los ajustes usamos el parámetro de mezcla $\alpha \approx \sqrt{2}/2$.

A.1. Caso Normal

Para este caso las entradas de la matriz simétrica X son variables aleatorias normales estándar con media cero y varianza 1.

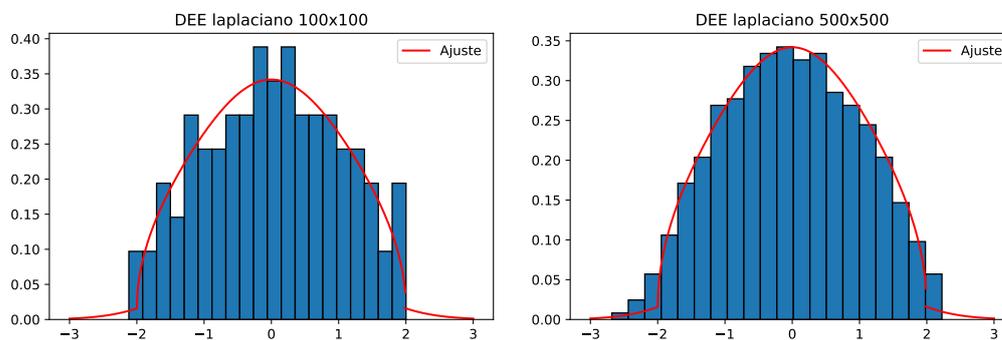
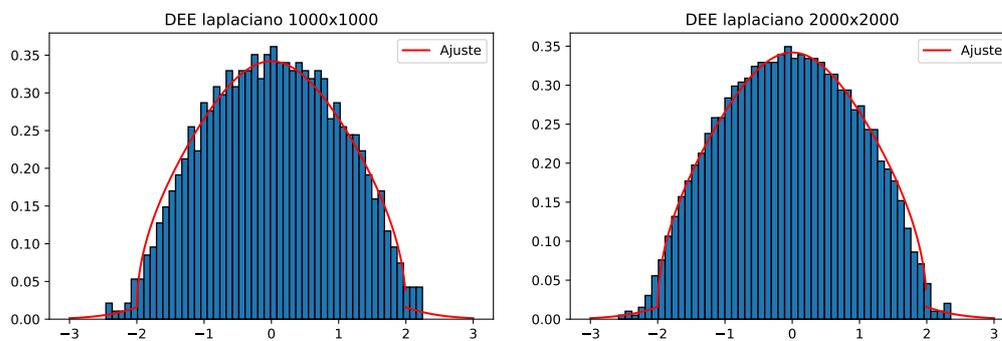
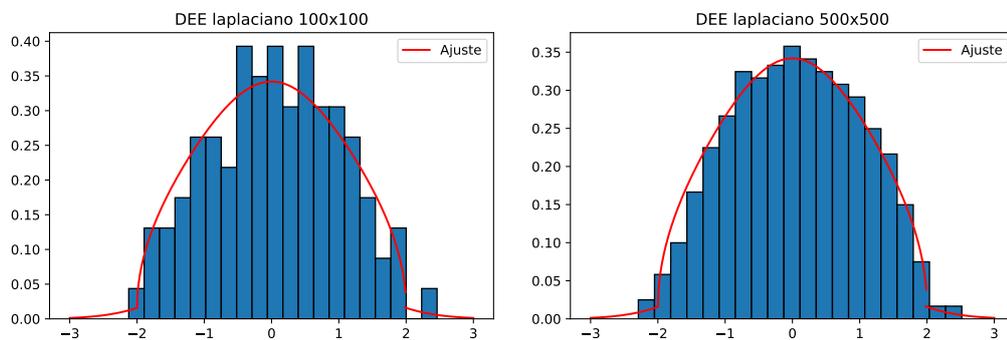
(a) Ajuste matriz laplaciana 100×100 .(b) Ajuste matriz laplaciana 500×500 .(c) Ajuste matriz laplaciana 1000×1000 .(d) Ajuste matriz laplaciana 2000×2000 .

Figura A.1: Histogramas de la distribución empírica espectral DEE y el ajuste dado por la conjetura para matrices laplacianas $L_{[X]}$ normalizadas por $\sqrt{2n}$ con las entradas de X normal estándar.

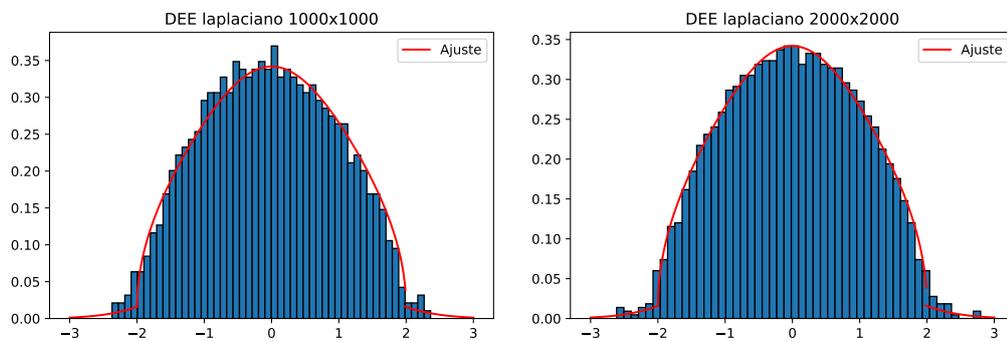
A.2. Caso uniforme

Necesitamos que las variables aleatorias tengan media cero y varianza uno, de esta manera, para este caso consideramos que las entradas de la matriz X están dadas por $Unif(-\sqrt{3}, \sqrt{3})$.



(a) Ajuste matriz laplaciana 100×100 .

(b) Ajuste matriz laplaciana 200×200 .



(c) Ajuste matriz laplaciana 1000×1000 .

(d) Ajuste matriz laplaciana 100×100 .

Figura A.2: Histogramas de la distribución empírica espectral DEE y el ajuste dado por la conjetura para matrices laplacianas $L_{[X]}$ normalizadas por $\sqrt{2n}$ con las entradas de X dadas por $Unif(-\sqrt{3}, \sqrt{3})$.

A.3. Caso Bernoulli

En este caso consideramos entradas fuera de la diagonal como variables aleatorias $Ber(1/2)$ tomando valores en $\{-1, 1\}$.

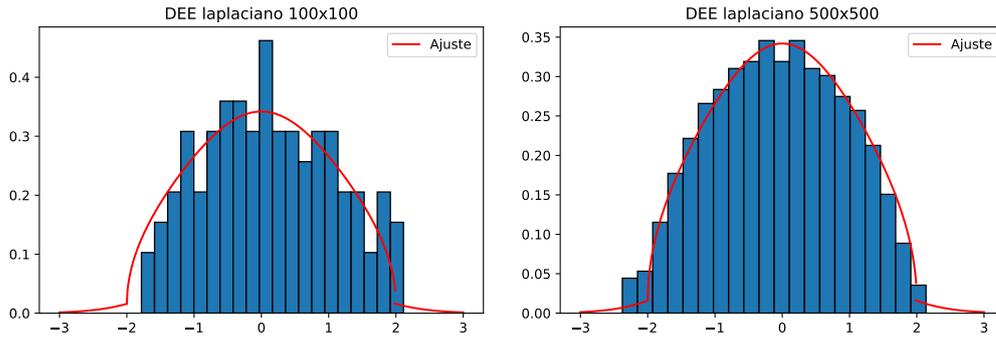
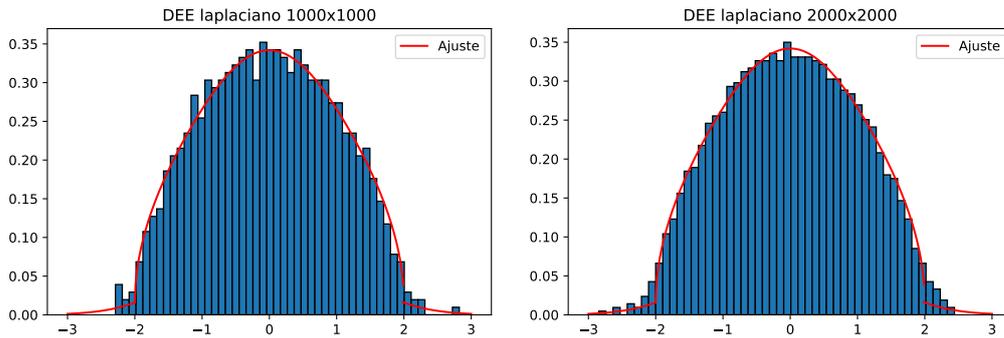
(a) Ajuste matriz laplaciana 100×100 .(b) Ajuste matriz laplaciana 500×500 .(c) Ajuste matriz laplaciana 1000×1000 .(d) Ajuste matriz laplaciana 2000×2000 .

Figura A.3: Histogramas de la distribución empírica espectral DEE y el ajuste dado por la conjetura para matrices laplacianas $L_{[X]}$ normalizadas por $\sqrt{2n}$ con las entradas de X dadas por $Ber(1/2)$ tomando valores en $\{-1, 1\}$.

Bibliografía

- [1] Emmanuel Abbe, Afonso S Bandeira, Annina Bracher, and Amit Singer. Decoding binary node labels from censored edge measurements: Phase transition and efficient recovery. *IEEE Transactions on Network Science and Engineering*, **1**(1):10–22, 2014.
- [2] Konstantin Avrachenkov, Laura Cottatellucci, and Arun Kadavankandy. Spectral properties of random matrices for stochastic block model. In *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), 2015 13th International Symposium on*, pages 537–544. IEEE, 2015.
- [3] Z.D Bai and Y.Q Yin. Limit of the smallest eigenvalue of a large dimensional sample covariance matrix. *The Annals of Probability*, **21**(3):1275–1294, 1993.
- [4] Zhi-Dong Bai and Yong-Qua Yin. Necessary and sufficient conditions for almost sure convergence of the largest eigenvalue of a wigner matrix. *The Annals of Probability*, **16**(4):1729–1741, 1988.
- [5] Zhidong D Bai. Methodologies in spectral analysis of large dimensional random matrices, a review. In *Advances In Statistics*, pages 174–240. World Scientific, 2008.
- [6] Afonso S Bandeira. Random laplacian matrices and convex relaxations. *Foundations of Computational Mathematics*, **18**:345–379, 2018.

- [7] Simeon M Berman. Limit theorems for the maximum term in stationary sequences. *The Annals of Mathematical Statistics*, **35**(2):502–516, 1964.
- [8] Włodzimierz Bryc, Amir Dembo, and Tiefeng Jiang. Spectral measure of large random hankel, markov and toeplitz matrices. *The Annals of Probability*, **34**:1–38, 2006.
- [9] Xue Ding and Tiefeng Jiang. Spectral distributions of adjacency and laplacian matrices of random graphs. *The Annals of Applied Probability*, **20**(6):2086–2117, 2010.
- [10] Paul Erdős and Alfréd Rényi. On random graphs, i. *Publicationes Mathematicae (Debrecen)*, **6**:290–297, 1959.
- [11] Paul Erdős and Alfréd Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, **5**(17-61):43, 1960.
- [12] Monique Laurent and Frank Vallentin. Semidefinite optimization. *Lecture Notes*, available at <http://page.mi.fu-berlin.de/fmario/sdp/laurentv.pdf>, 2012.
- [13] Sidney I Resnick. Extreme values, regular variation, and point processes, volume 4 of applied probability. a series of the applied probability trust, 1987.
- [14] Eugene P Wigner. On the distribution of the roots of certain symmetric matrices. *Annals of Mathematics*, **67**:325–327, 1958.