

El teorema de Wigner para matrices aleatorias¹

J. Armando Domínguez Molina

jadguez@uas.uasnet.mx

Alfonso Rocha Arteaga

arteaga@uas.uasnet.mx

Facultad de Ciencias Físico-Matemáticas
Universidad Autónoma de Sinaloa

Resumen

El propósito de este artículo es dar una demostración detallada del teorema de Wigner, el cual es un resultado central de la teoría de matrices aleatorias, usando herramientas de teoría de gráficas, combinatoria y probabilidad.

1. Introducción

El teorema de Wigner nos dice que la distribución espectral empírica de ciertas matrices aleatorias converge a la distribución del semicírculo cuando la dimensión de éstas tiende a infinito. Existen varias demostraciones de este resultado básico de la teoría de matrices aleatorias, sin embargo en este artículo se pretende exponer una demostración detallada usando herramientas elementales de teoría de gráficas, combinatoria y probabilidad.

Las variables aleatorias que aquí consideraremos estarán definidas en un espacio de probabilidad $(\Omega, \mathfrak{F}, \mathbf{P})$ donde Ω es el espacio muestral y \mathbf{P} es la medida de probabilidad definida sobre la σ -álgebra de eventos \mathfrak{F} .

Una *matriz aleatoria* es una matriz cuyas entradas son variables aleatorias. Llamaremos *matriz aleatoria hermitiana* a una matriz aleatoria con entradas complejas que coincide con su transpuesta conjugada

¹Existe una versión preliminar de este trabajo en Comunicaciones del CIMAT, I-09-08 (PE). PROFAPI 2010-044

(el caso real corresponde a las matrices aleatorias simétricas). Una propiedad importante de las matrices aleatorias hermitianas es que sus valores propios son reales y son aleatorios.

Sea A_N una matriz aleatoria hermitiana de $N \times N$ y sean las variables aleatorias reales $\lambda_1^N, \lambda_2^N, \dots, \lambda_N^N$ sus valores propios. La *función de distribución espectral empírica* de A_N ,

$$F_{A_N}(x) = \frac{1}{N} \# \{i : \lambda_i^N \leq x\} \quad x \in \mathbb{R},$$

determina la proporción de valores propios menores o iguales que x . El término espectral se usa para hacer referencia a los valores propios. Ésta es una función de distribución aleatoria, por lo tanto para cada ω del espacio muestral $F_{A_N(\omega)}$ es una función de distribución y es precisamente la función de distribución empírica correspondiente a los números reales $\lambda_1^N(\omega), \lambda_2^N(\omega), \dots, \lambda_N^N(\omega)$. La función de distribución espectral empírica se puede interpretar como una función de distribución aleatoria uniforme en los valores propios $\lambda_1^N, \lambda_2^N, \dots, \lambda_N^N$, donde cada uno de estos valores propios tiene la misma probabilidad $\frac{1}{N}$ de ocurrir.

El teorema de Wigner establece, bajo ciertas condiciones sobre A_N , que para cada $x \in \mathbb{R}$ cuando $N \rightarrow \infty$

$$F_{A_N}(x) \xrightarrow{c.s.} F(x) = \int_{-\infty}^x \frac{1}{2\pi} \sqrt{4 - t^2} 1_{[-2,2]}(t) dt, \quad (1)$$

donde $F(x)$ es la *función de distribución del semicírculo* cuya función de densidad está dada por $f(t) = \frac{1}{2\pi} \sqrt{4 - t^2}$, $-2 \leq t \leq 2$. Esta convergencia es la convergencia casi segura, es decir salvo en un conjunto de probabilidad cero, se cumple que para cada ω del espacio muestral la distribución espectral empírica converge puntualmente a la distribución del semicírculo.

Este resultado lo obtuvo E. P. Wigner [11] en la década de los cincuenta para el caso de matrices aleatorias simétricas con entradas reales, independientes (salvo simetría), simétricamente distribuidas, centradas, con varianza constante y con momentos de cualquier orden uniformemente acotados. Su trabajo estuvo motivado en mecánica cuántica, donde logró explicar el comportamiento estadístico de los niveles de energía de un sistema físico en términos de los valores propios cuando la dimensión de las matrices es grande.

Existen varias versiones de este resultado dependiendo de las condiciones que se consideren sobre las matrices aleatorias, [1], [2], [4], [5], [6], [8] y [9].

En este trabajo se demostrará la convergencia (1) por el método de momentos. Concretamente, una función de distribución F tiene *momento de orden k* si $\int |x|^k F(dx)$ existe. La función de distribución F está *únicamente determinada por sus momentos* si existen sus momentos de orden k , $k = 1, 2, \dots$, y si G es cualquier función de distribución tal que $\int x^k F(dx) = \int x^k G(dx)$ entonces $F = G$. El método de momentos establece lo siguiente: sean F_N y F funciones de distribución para las cuales existen sus momentos de orden k , para $k = 1, 2, \dots$, respectivamente. Si F está únicamente determinada por sus momentos y $\int x^k F_N(dx) \rightarrow \int x^k F(dx)$, cuando $N \rightarrow \infty$, para $k = 1, 2, \dots$, entonces $F_N(x) \rightarrow F(x)$ cuando $N \rightarrow \infty$ para cada punto de continuidad de F .

Por un lado, los momentos de la función de distribución espectral empírica cumplen la relación

$$\int x^k F_{A_N}(dx) = \frac{\text{tr}(A_N^k)}{N} \quad c.s. \quad k = 0, 1, 2, \dots, \quad (2)$$

donde $\text{tr}(A_N^k)$ es la traza de la potencia k de la matriz A_N , ya que

$$\begin{aligned} \int x^k F_{A_N}(dx) &= \sum_{j=1}^N (\lambda_j^N)^k [F_{A_N}(\lambda_j^N) - F_{A_N}(\lambda_j^N -)] \\ &= \frac{1}{N} \sum_{j=1}^N (\lambda_j^N)^k = \frac{\text{tr}(A_N^k)}{N}. \end{aligned}$$

Por otra parte, como veremos más adelante, la distribución del semicírculo está únicamente determinada por sus momentos

$$\int x^k F(dx) = C_{\frac{k}{2}} \quad \text{si } k \text{ par} \quad \text{o} \quad 0 \text{ si } k \text{ impar}, \quad (3)$$

donde C_k son los *números de Catalan* los cuales están definidos como

$$C_k := \frac{1}{k+1} \binom{2k}{k} \quad k = 0, 1, 2, \dots \quad (4)$$

Por lo tanto, de acuerdo al método de momentos, para demostrar la convergencia (1) es suficiente demostrar la convergencia casi segura de (2) a (3), esto es

$$\frac{\text{tr}(A_N^k)}{N} \xrightarrow{c.s.} C_{\frac{k}{2}} \quad \text{si } k \text{ par} \quad \text{o} \quad 0 \text{ si } k \text{ impar}. \quad (5)$$

Más aún, probaremos la convergencia en esperanza

$$\mathbf{E} \left[\frac{\text{tr}(A_N^k)}{N} \right] \longrightarrow C_{\frac{k}{2}} \quad \text{si } k \text{ par} \quad \text{o} \quad 0 \text{ si } k \text{ impar} \quad (6)$$

y usaremos este resultado para probar (5). Esta última convergencia la demostraremos usando el enfoque combinatorio de gráficas. Aquí la clave es que la traza presenta una estructura cíclica que permite asociarle gráficas conexas, de manera que al tomar el límite en (6), quedarán sólo aquellas gráficas conexas correspondientes a los árboles orientados con raíz, los cuales vienen dados por los números de Catalan.

Las condiciones bajo las cuales se probará (6), y por lo tanto el teorema de Wigner, son las siguientes: la matriz aleatoria A_N es compleja de $N \times N$, hermitiana $(A_N)_{ij} = \overline{(A_N)_{ji}}$, con entradas independientes $(A_N)_{ij}$ $1 \leq i \leq j \leq N$, centradas $\mathbf{E}(A_N)_{ij} = 0$, con varianza $\mathbf{E}|(A_N)_{ij}|^2 = 1/N$ y momentos de todos los órdenes $\mathbf{E}|\sqrt{N}(A_N)_{ij}|^k$ uniformemente acotados en i, j y N por una constante que sólo depende de k .

Estas condiciones son esencialmente las mismas que considera Wigner [11] para matrices aleatorias simétricas reales, salvo por la distribución simétrica de las entradas.

Este artículo está organizado como sigue. En la sección 2 se revisan algunos conceptos básicos de probabilidad y se ve que la distribución del semicírculo está únicamente determinada por sus momentos (3) y (4), los cuales están dados por los números de Catalan. En la sección 3 se revisan algunos conceptos básicos de teoría de gráficas, poniendo especial interés en los árboles. Se ve que la cantidad de árboles orientados con raíz viene dada por los números de Catalan. Este resultado se utiliza en la sección 4 para probar la convergencia en esperanza (6) y ésta se usa para probar la convergencia casi segura (5). Finalmente en la sección 5 se comentan algunas variantes del teorema de Wigner.

2. Elementos de probabilidad

En esta sección veremos los elementos de probabilidad que justifican el método empleado para la demostración del resultado principal mencionado en la introducción. Particularmente, recordaremos el método de momentos y que la función de distribución del semicírculo está únicamente determinada por sus momentos, los cuales vienen dados por los números de Catalan.

2.1. Momentos

Sea $(\Omega, \mathfrak{F}, \mathbf{P})$ un espacio de probabilidad donde Ω es el espacio muestral y \mathbf{P} es la medida de probabilidad definida sobre la σ -álgebra \mathfrak{F} de subconjuntos de Ω . A los elementos de \mathfrak{F} se les llama *eventos*. Una *variable aleatoria* es una función $X : \Omega \rightarrow \mathbb{R}$ Borel medible, es decir, para todo conjunto de Borel B en \mathbb{R} el conjunto $\{\omega \in \Omega : X(\omega) \in B\}$ pertenece a \mathfrak{F} . En adelante denotaremos a estos eventos definidos por la variable aleatoria como $\{X \in B\}$ o $X \in B$.

La *función de distribución* de una variable aleatoria X es la función $F : \mathbb{R} \rightarrow [0, 1]$ definida por $F(x) = \mathbf{P}(X \leq x)$ para cada x en \mathbb{R} . Ésta determina el comportamiento probabilístico de la variable aleatoria, donde $F(x)$ se interpreta como la probabilidad de que la variable aleatoria sea menor o igual a x .

Las funciones de distribución están caracterizadas de la siguiente manera. Una función $F : \mathbb{R} \rightarrow [0, 1]$ es la función de distribución de una variable aleatoria X definida en algún espacio de probabilidad si y sólo si cumple las siguientes tres propiedades: (i) F es continua por la derecha, (ii) F es no decreciente y (iii) $F(x) \rightarrow 0$ cuando $x \rightarrow -\infty$ y $F(x) \rightarrow 1$ cuando $x \rightarrow \infty$.

El *momento de orden k* de una variable aleatoria X (o de su función de distribución F) se define, en caso de existir, como el número

$$\mathbf{E}X^k := \int_{\mathbb{R}} x^k F(dx),$$

donde la integral se toma en el sentido de Stieltjes y es absolutamente convergente, ver capítulo 2 de [10]. La *esperanza* de X se define, en caso de existir el momento de orden 1, como $\mathbf{E}X$. La *varianza* de X se define, en caso de existir momento de orden 2, como $\text{var}(X) =: \mathbf{E}[X - \mathbf{E}X]^2 = \mathbf{E}X^2 - [\mathbf{E}X]^2 \geq 0$.

La desigualdad de Chebyshev establece una relación importante entre la esperanza y la varianza de cualquier variable aleatoria, ver [10].

Proposición 1 (Desigualdad de Chebyshev) *Sea X una variable aleatoria con varianza finita. Entonces para cualquier constante $\varepsilon > 0$,*

$$\mathbf{P}(|X - \mathbf{E}X| \geq \varepsilon) \leq \frac{\text{var}(X)}{\varepsilon^2}.$$

Si A_1, \dots, A_n son eventos tales que $\mathbf{P}(A_1 \cap \dots \cap A_n) = \mathbf{P}(A_1) \cdots \mathbf{P}(A_n)$ entonces se llaman *eventos independientes*. Si X_1, \dots, X_n son variables aleatorias tales que los eventos $\{X_1 \leq x_1\}, \dots, \{X_n \leq x_n\}$ son independientes para cualquier elección de $x_1 \in \mathbb{R}, \dots, x_n \in \mathbb{R}$ entonces se llaman

variables aleatorias independientes. Si X_1, \dots, X_n son variables aleatorias independientes cuyas esperanzas existen entonces $\mathbf{E}[X_1 \cdots X_n] = \mathbf{E}X_1 \cdots \mathbf{E}X_n$.

2.2. Método de momentos

Sean X, X_1, X_2, \dots , variables aleatorias con funciones de distribución F, F_1, F_2, \dots , respectivamente.

Se dice que X_n *converge casi seguramente* a X , o que $X_n \xrightarrow{c.s.} X$, si existe un evento Ω_0 tal que $P(\Omega_0) = 1$ y $X_n(\omega) \rightarrow X(\omega)$ cuando $n \rightarrow \infty$ para cada $\omega \in \Omega_0$.

Si *los momentos de X_n convergen a los momentos de X* , esto es $\mathbf{E}X_n^k \rightarrow \mathbf{E}X^k$ cuando $n \rightarrow \infty$ para cada $k \geq 1$, no necesariamente $X_n \xrightarrow{c.s.} X$. Sin embargo, en la sección 4.3 demostraremos la convergencia de los momentos (6) y la usaremos en la sección 4.4 para demostrar la convergencia casi segura (5).

La convergencia de los momentos tampoco implica la *convergencia en distribución* de las respectivas funciones de distribución, es decir la convergencia $F_n(x) \rightarrow F(x)$ cuando $n \rightarrow \infty$ para cada x punto de continuidad de F . El siguiente resultado conocido como el método de momentos, proporciona un criterio bajo el cual la convergencia de los momentos implica la convergencia en distribución. Para una justificación de este método, ver el apéndice *B* de [2].

Una función de distribución F de una variable aleatoria X está *únicamente determinada por sus momentos* si existen sus momentos de orden k , para $k = 1, 2, \dots$, y si G es cualquier función de distribución de una variable aleatoria Y tal que $\mathbf{E}X^k = \mathbf{E}Y^k$ entonces $F = G$.

Teorema 2 (Método de momentos) *Sean F_n y F funciones de distribución de las variables aleatorias X_n y X respectivamente, para las cuales existen todos sus momentos de orden $k \geq 1$. Si F está determinada por sus momentos, y si para cada k , el k -ésimo momento $\mathbf{E}X_n^k \rightarrow \mathbf{E}X^k$ cuando $n \rightarrow \infty$ entonces $F_n(x) \rightarrow F(x)$ para cada x punto de continuidad de F .*

La función de distribución del semicírculo $F(x)$ en (1) está únicamente determinada por sus momentos. Esto se sigue de que tiene soporte acotado y sus momentos están dados por los números de Catalan, ver el apéndice *B* de [2]. Enunciamos ahora la expresión de sus momentos, cuya demostración se puede ver en la sección 2.1.1 de [2].

Lema 3 *Los momentos pares de la distribución del semicírculo están dados por los números de Catalan*

$$\int_{\mathbb{R}} t^{2k} \frac{1}{2\pi} \sqrt{4-t^2} 1_{[-2,2]}(t) dt = \frac{1}{k+1} \binom{2k}{k}.$$

Los momentos impares son cero debido a la simetría de la distribución.

3. Elementos de teoría de gráficas

En esta sección veremos algunos elementos de combinatoria y teoría de gráficas que emplearemos como técnica para demostrar el teorema de Wigner de la sección 4. En particular recordaremos que los números de Catalan enumeran los árboles orientados con raíz.

3.1. Gráficas

Una *gráfica* es un par ordenado $(V(G), E(G))$ que consiste de un conjunto $V(G)$ de *vértices* y un conjunto $E(G)$, disjunto de $V(G)$, de *aristas*, junto con una función ψ_G que asocia a cada arista de G un par no ordenado de vértices de G no necesariamente distintos. Si $\psi_G(e) = \{u, v\}$ diremos que los vértices u y v son los *extremos* de la arista e . El número de vértices y el número de aristas de G los denotaremos por $|V(G)|$ y $|E(G)|$, respectivamente.

Un *lazo* es una arista con extremos idénticos. Dos *aristas paralelas* son dos aristas que tienen los mismos extremos. Una *gráfica simple* es una gráfica que no tiene lazos ni aristas paralelas.

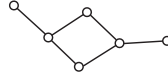
Un *ciclo con tres o más vértices*, es una gráfica simple cuyos vértices se pueden disponer en una secuencia cíclica de tal manera que dos vértices son los extremos de una misma arista si ellos son consecutivos en la secuencia y no son los extremos de una misma arista si no son consecutivos. Un *ciclo con un vértice* consiste de un vértice con un lazo. Un *ciclo con dos vértices* consiste en dos vértices que son los extremos de dos aristas paralelas.

Una *gráfica conexa*, es una gráfica tal que para toda partición de su conjunto de vértices en dos conjuntos no vacíos, existe una arista con un extremo en uno de esos conjuntos y el otro extremo en el otro conjunto. Una *gráfica inconexa* es una gráfica que no es conexa.

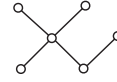
Una gráfica F es una *subgráfica* de una gráfica G si $V(F) \subseteq V(G)$, $E(F) \subseteq E(G)$ y la función ψ_F es la restricción de ψ_G a $E(F)$. Se dice que G contiene a F . Si la subgráfica F tiene el mismo conjunto de



gráfica (inconexa)



gráfica conexa (con ciclo)



árbol

vértices que G , esto es $V(F) = V(G)$, diremos que F es una *subgráfica recubridora* de G .

Un *árbol* es una gráfica conexa que no contiene ciclos. Una subgráfica recubridora que es árbol se llama *árbol recubridor*. Una gráfica es conexa si y sólo si tiene un árbol recubridor, ver [3].

Nos interesa la relación entre el número de vértices y aristas de las gráficas conexas. Sea G una gráfica conexa. Si G es un árbol entonces $|V(G)| = |E(G)| + 1$ y recíprocamente, ver sección 4.1 de [3]. Como toda gráfica conexa tiene un árbol recubridor se deduce de aquí el siguiente resultado, que nos dice que las gráficas conexas no pueden tener más vértices que aristas; salvo el caso de los árboles donde el número de vértices supera en uno al número de aristas.

Lema 4 *Si $G = (V(G), E(G))$ es una gráfica conexa entonces*

$$|V(G)| \leq |E(G)| + 1.$$

Además, $|V(G)| = |E(G)| + 1$ si y sólo si G es árbol.

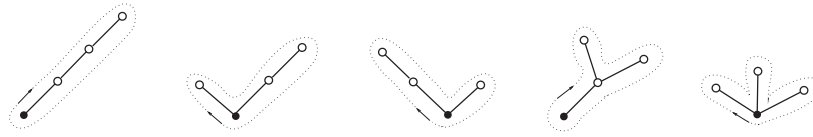
Un *árbol con raíz* es un árbol al que se le ha especificado un vértice al que llamaremos *raíz*. Un *árbol orientado con raíz* es un árbol con raíz inmerso en un plano, cuyo contorno se recorre sobre el plano siempre en un mismo sentido previamente convenido, de la siguiente manera: el recorrido inicia y termina en la raíz, al término del recorrido se recorre la misma cantidad de aristas de ida que de vuelta y en cada paso del recorrido no se tienen más aristas recorridas de vuelta que de ida.

Nos interesa contar *el número de árboles orientados con raíz que se pueden formar con k aristas*. Existen en la literatura varias maneras de contarlos. Una de ellas es identificándolos con $2k$ -tuplas de entradas dicotómicas como sigue. A cada árbol orientado con raíz de k aristas se le asocia la $2k$ -tupla (ξ_1, \dots, ξ_{2k}) construida de la siguiente manera: si en el i -ésimo paso del recorrido la arista en turno es recorrida por primera vez entonces $\xi_i = 1$ y si ya se había pasado ella entonces $\xi_i = -1$. Esto establece una biyección entre los árboles orientados con raíz de k aristas y el subconjunto T_{2k} de $\{-1, 1\}^{2k}$ que consiste de los elementos

(ξ_1, \dots, ξ_{2k}) tales que

$$\begin{aligned} \xi_i &= \pm 1 \\ \xi_1 + \xi_2 + \dots + \xi_{2k} &= 0 \\ \xi_1 + \xi_2 + \dots + \xi_j &\geq 0 \quad j = 1, 2, \dots, 2k. \end{aligned}$$

Ejemplo 5 ($k = 3$) *Con tres aristas se forman cinco árboles orientados con raíz, los cuales corresponden a las 6-tuplas $(1, 1, 1, -1, -1, -1)$, $(1, -1, 1, 1, -1, -1)$, $(1, 1, -1, -1, 1, -1)$, $(1, 1, -1, 1, -1, -1)$, $(1, -1, 1, -1, 1, -1)$, respectivamente.*



El siguiente resultado se sigue de que la cardinalidad de T_{2k} es el k -ésimo número de Catalan, ver capítulo 6 de [7].

Proposición 6 *El número de árboles orientados con raíz que se pueden formar con k aristas (o equivalentemente $k + 1$ vértices) coincide con el k -ésimo número de Catalan*

$$C_k = \frac{1}{k + 1} \binom{2k}{k} \quad k \geq 1.$$

En otras palabras, los números de Catalan cuentan el total de árboles orientados con raíz que se pueden formar con un número fijo de aristas.

3.2. Gráficas dirigidas

Una *gráfica dirigida* es un par ordenado $(V(G), E(G))$ que consiste de un conjunto $V(G)$ de *vértices* y un conjunto $E(G)$, disjunto de $V(G)$, de *aristas dirigidas*, junto con una función ψ_G que asocia a cada arista dirigida de G un par ordenado de vértices de G no necesariamente distintos. Si $\psi_G(e) = (u, v)$ diremos que los vértices u y v son *los extremos de la arista dirigida e* y que ésta va del extremo u al extremo v . Por simplicidad, abreviaremos el término “gráfica dirigida” por *digráfica* y “arista dirigida” por *diarista*.

Una diarista con extremos idénticos se llama *lazo*. Dos *diaristas paralelas* son dos diaristas que tienen los mismos extremos. El ejemplo

10 muestra dos representaciones distintas de una digráfica con un lazo y tres diaristas paralelas.

A cada digráfica G se le asocia una gráfica con el mismo conjunto de vértices simplemente reemplazando cada diarista por una arista con los mismos extremos, a esta gráfica se le llama la *gráfica subyacente* de la digráfica G . Observemos que una digráfica es conexa si y sólo si su gráfica subyacente es conexa.

El *esqueleto* de una digráfica G es la gráfica que se obtiene de la gráfica subyacente de G al eliminar las aristas paralelas. Las digráficas conexas tienen esqueletos conexos. El ejemplo 10 muestra una digráfica conexa con su respectivo esqueleto conexo.

4. El teorema de Wigner

Una *matriz de Wigner* es una matriz aleatoria A^N de $N \times N$ con elementos reales o complejos tales que A_{ij}^N $1 \leq i \leq j \leq N$ son independientes y A^N hermitiana $A_{ij}^N = \overline{A_{ji}^N}$.

Consideraremos en adelante sólo matrices de Wigner que cumplen las siguientes tres condiciones:

$$\mathbf{E}A_{ij}^N = 0, \quad (7a)$$

$$\mathbf{E}|A_{ij}^N|^2 = 1/N, \quad (7b)$$

$$M_k := \sup_{N \in \mathbb{N}} \max_{1 \leq i, j \leq N} \mathbf{E} \left| \sqrt{N} A_{ij}^N \right|^k < \infty. \quad (7c)$$

4.1. Resultados

Teorema 7 (Teorema de Wigner) *Si A^N , $N \geq 1$, es una sucesión de matrices de Wigner que cumplen las tres condiciones en (7) entonces la distribución espectral empírica de A^N converge casi seguramente a la distribución del semicírculo, es decir, para cada $x \in \mathbb{R}$, cuando $N \rightarrow \infty$*

$$F_{A^N}(x) := \frac{1}{N} \# \{i : \lambda_i^N \leq x\} \xrightarrow{c.s.} \int_{-\infty}^x \frac{1}{2\pi} \sqrt{4 - t^2} 1_{[-2,2]}(t) dt.$$

Como se mencionó en la introducción, para demostrar el teorema de Wigner es suficiente de acuerdo al método de momentos demostrar los siguientes dos resultados.

Teorema 8 *Bajo las mismas condiciones del teorema 7 sobre las matrices aleatorias A^N se cumple la convergencia en esperanza*

$$\mathbf{E} \left\{ \frac{\text{tr} \left[(A^N)^k \right]}{N} \right\} \longrightarrow \begin{cases} C_{\frac{k}{2}} & \text{si } k \text{ es par,} \\ 0 & \text{si } k \text{ es impar.} \end{cases}$$

Teorema 9 *Bajo las mismas condiciones del teorema 7 sobre las matrices aleatorias A^N se cumple la convergencia casi segura*

$$\frac{\text{tr} \left[(A^N)^k \right]}{N} \xrightarrow{\text{c.s.}} \begin{cases} C_{\frac{k}{2}} & \text{si } k \text{ es par,} \\ 0 & \text{si } k \text{ es impar.} \end{cases}$$

4.2. Traza y gráficas conexas

Para las demostraciones de los teoremas 7, 8 y 9, introduciremos cierta notación y haremos algunas consideraciones preliminares acerca de la traza de potencias de matrices.

Sean

$$B^N := \sqrt{N}A^N \quad \text{y} \quad B_{ij} := \sqrt{N}A_{ij}^N.$$

$$\mathbf{i} := (i_1, \dots, i_k) \in \{1, 2, \dots, N\}^k, \quad i_{k+1} := i_1.$$

$$Q(\mathbf{i}) := \mathbf{E} [B_{i_1 i_2} B_{i_2 i_3} \cdots B_{i_k i_1}].$$

$$\sum_{\mathbf{i}} := \sum_{i_1, \dots, i_k=1}^N := \sum_{i_1=1}^N \cdots \sum_{i_k=1}^N.$$

Teniendo en cuenta la regla $(CD)_{ij} = \sum_{k=1}^m C_{ik} D_{kj}$ del producto de dos matrices m -compatibles C y D , podemos escribir la traza de la k -ésima potencia de B^N como

$$\begin{aligned} \text{tr} \left((B^N)^k \right) &= \sum_{i_1=1}^N \left((B^N)^k \right)_{i_1 i_1} = \sum_{i_1, i_2=1}^N B_{i_1 i_2} \left((B^N)^{k-1} \right)_{i_2 i_1} \\ &= \cdots = \sum_{i_1, \dots, i_k=1}^N B_{i_1 i_2} B_{i_2 i_3} \cdots B_{i_k i_1}, \end{aligned}$$

de donde

$$\begin{aligned}
 \mathbf{E} \left[\frac{1}{N} \operatorname{tr} \left[(A^N)^k \right] \right] &= \mathbf{E} \left[\frac{1}{N} \operatorname{tr} \left[\left(N^{-\frac{1}{2}} B^N \right)^k \right] \right] \\
 &= \frac{1}{N^{\frac{k}{2}+1}} \sum_{i_1, \dots, i_k=1}^N \mathbf{E} [B_{i_1 i_2} B_{i_2 i_3} \cdots B_{i_k i_1}] \\
 &= \frac{1}{N^{\frac{k}{2}+1}} \sum_{\mathbf{i}} Q(\mathbf{i}). \tag{7}
 \end{aligned}$$

Vemos de aquí que la secuencia cíclica que la traza induce a $Q(\mathbf{i})$ permite identificar a cada índice $\mathbf{i} = (i_1, \dots, i_k)$ de la suma anterior con una digráfica conexa $G(\mathbf{i}) = (V(\mathbf{i}), E(\mathbf{i}))$ definida como

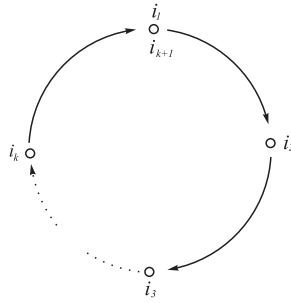
$$\begin{aligned}
 V(\mathbf{i}) &= \{i_1, \dots, i_k\} \text{ vértices,} \\
 E(\mathbf{i}) &= \{e_1, \dots, e_k\} \text{ diaristas,}
 \end{aligned}$$

$$\psi_{G(\mathbf{i})}(e_1) = (i_1, i_2), \dots, \psi_{G(\mathbf{i})}(e_{k-1}) = (i_{k-1}, i_k), \psi_{G(\mathbf{i})}(e_k) = (i_k, i_1).$$

En adelante denotaremos por simplicidad a este conjunto de diaristas por

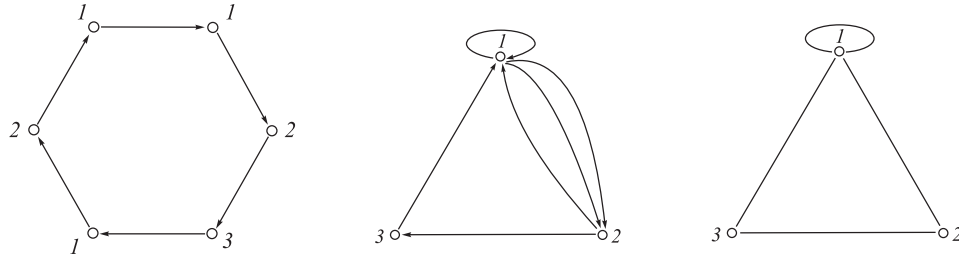
$$E(\mathbf{i}) = \{(i_1, i_2), \dots, (i_{k-1}, i_k), (i_k, i_1)\}.$$

Esta digráfica conexa se puede representar por la siguiente figura que ilustra la secuencia cíclica.



A cada una de estas digráficas conexas $G(\mathbf{i})$ le asociamos su esqueleto $\tilde{G}(\mathbf{i}) = (\tilde{V}(\mathbf{i}), \tilde{E}(\mathbf{i}))$, el cual es la gráfica subyacente de $G(\mathbf{i})$ sin aristas paralelas.

Ejemplo 10 Si $G(\mathbf{i}) = (V(\mathbf{i}), E(\mathbf{i}))$ con $V(\mathbf{i}) = \mathbf{i} = \{1, 2, 3, 1, 2, 1\}$ y $E(\mathbf{i}) = \{(1, 2), (2, 3), (3, 1), (1, 2), (2, 1), (1, 1)\}$ entonces su esqueleto es $\tilde{G}(\mathbf{i}) = (\tilde{V}(\mathbf{i}), \tilde{E}(\mathbf{i}))$ donde $\tilde{V}(\mathbf{i}) = \{1, 2, 3\}$ y $\tilde{E}(\mathbf{i}) = \{\{1, 2\}, \{2, 3\}, \{3, 1\}, \{1, 1\}\}$. Las dos primeras figuras corresponden a distintas representaciones de la digráfica $G(\mathbf{i})$ mientras que la tercera es una representación de su esqueleto $\tilde{G}(\mathbf{i})$.



Podemos descomponer la suma (7) clasificando sus términos de acuerdo al número de vértices que tienen los esqueletos $\tilde{G}(\mathbf{i})$ correspondientes a estas digráficas $G(\mathbf{i})$

$$\begin{aligned} \frac{1}{N^{\frac{k}{2}+1}} \sum_{\mathbf{i}} Q(\mathbf{i}) &= \frac{1}{N^{\frac{k}{2}+1}} \sum_{G(\mathbf{i})} Q(\mathbf{i}) \\ &= \frac{1}{N^{\frac{k}{2}+1}} \sum_{\substack{G(\mathbf{i}) \\ 1 \leq |\tilde{V}(\mathbf{i})| \leq \lfloor \frac{k}{2} \rfloor + 1}} Q(\mathbf{i}) + \frac{1}{N^{\frac{k}{2}+1}} \sum_{\substack{G(\mathbf{i}) \\ |\tilde{V}(\mathbf{i})| \geq \lfloor \frac{k}{2} \rfloor + 2}} Q(\mathbf{i}). \end{aligned}$$

Los dos lemas siguientes muestran que esta última suma es cero, ya que sus correspondientes $Q(\mathbf{i})$ no contribuyen a ella. Antes observemos que, dos aristas de $E(\mathbf{i})$ no son paralelas si no tienen los mismos extremos. En el ejemplo 10 las diaristas $(2, 3)$, $(3, 1)$ y $(1, 1)$ no tienen diaristas paralelas en $E(\mathbf{i})$ mientras que $(1, 2)$ y $(2, 1)$ son diaristas paralelas.

Lema 11 Si $E(\mathbf{i})$ tiene una diarista que no es paralela a ninguna otra diarista entonces $Q(\mathbf{i}) = 0$.

Demostración. De la definición de B^N es claro que $\mathbf{E}B_{ij} = 0$ para cada i, j y que B_{ij} y B_{lm} son independientes cuando $(i, j) \neq (l, m)$ y $(i, j) \neq (m, l)$. Luego si (i_l, i_{l+1}) es una diarista de $E(\mathbf{i})$ que no es paralela a ninguna otra diarista,

$$\begin{aligned} Q(\mathbf{i}) &= \mathbf{E}(B_{i_1 i_2} \cdots B_{i_{l-1} i_l} B_{i_l i_{l+1}} B_{i_{l+1} i_{l+2}} \cdots B_{i_k i_1}) \\ &= \mathbf{E}(B_{i_l i_{l+1}}) \mathbf{E}(B_{i_1 i_2} \cdots B_{i_{l-1} i_l} B_{i_{l+1} i_{l+2}} \cdots B_{i_k i_1}) = 0. \end{aligned}$$

■

Lema 12 Si $|\tilde{V}(\mathbf{i})| \geq \lfloor \frac{k}{2} \rfloor + 2$ entonces $E(\mathbf{i})$ tiene por lo menos una diarista que no es paralela a ninguna otra diarista.

Demostración. De acuerdo al lema 4 $|\tilde{E}(\mathbf{i})| + 1 \geq |\tilde{V}(\mathbf{i})|$ de donde $|\tilde{E}(\mathbf{i})| \geq \lfloor \frac{k}{2} \rfloor + 1$. Dado que se tienen $|E(\mathbf{i})| = k$ diaristas y hay por lo menos $\lfloor \frac{k}{2} \rfloor + 1$ aristas distintas, entonces $E(\mathbf{i})$ debe tener una diarista que no es paralela a ninguna otra diarista, ya que de lo contrario $E(\mathbf{i})$ tendría por lo menos $2(\lfloor \frac{k}{2} \rfloor + 1) = k + 2$ diaristas cuando k es par o por lo menos $k + 1$ diaristas cuando k es impar, lo cual en ambos casos no es posible. ■

Como consecuencia de los lemas 11 y 12 la expresión (7) se reduce a

$$\mathbf{E} \left[\frac{1}{N} \operatorname{tr} \left[(A^N)^k \right] \right] = \frac{1}{N^{\frac{k}{2}+1}} \sum_{\substack{G(\mathbf{i}) \\ 1 \leq |\tilde{V}(\mathbf{i})| \leq \lfloor \frac{k}{2} \rfloor + 1}} Q(\mathbf{i}). \quad (8)$$

Por otra parte, se tiene una cota uniforme para los $Q(\mathbf{i})$ al usar la desigualdad de Hölder y (7c)

$$\begin{aligned} |Q(\mathbf{i})| &= |\mathbf{E} [B_{i_1 i_2} B_{i_2 i_3} \cdots B_{i_k i_1}]| \leq \left[\mathbf{E} |B_{i_1 i_2}|^k \right]^{\frac{1}{k}} \cdots \left[\mathbf{E} |B_{i_k i_1}|^k \right]^{\frac{1}{k}} \\ &\leq M_k^{\frac{1}{k}} \cdots M_k^{\frac{1}{k}} = M_k. \end{aligned}$$

4.3. La convergencia en esperanza

Para la demostración de la convergencia en esperanza del teorema 8 usaremos los resultados preliminares de las secciones anteriores.

Demostración del teorema 8. Partiendo de (8)

$$\begin{aligned} \left| \mathbf{E} \left[\frac{1}{N} \operatorname{tr} \left[(A^N)^k \right] \right] \right| &\leq \frac{1}{N^{\frac{k}{2}+1}} \sum_{\substack{G(\mathbf{i}) \\ 1 \leq |\tilde{V}(\mathbf{i})| \leq \lfloor \frac{k}{2} \rfloor + 1}} |Q(\mathbf{i})| \leq \frac{1}{N^{\frac{k}{2}+1}} \sum_{\substack{G(\mathbf{i}) \\ 1 \leq |\tilde{V}(\mathbf{i})| \leq \lfloor \frac{k}{2} \rfloor + 1}} M_k \\ &\leq \frac{1}{N^{\frac{k}{2}+1}} \sum_{j=1}^{\lfloor \frac{k}{2} \rfloor + 1} \sum_{\substack{G(\mathbf{i}) \\ |\tilde{V}(\mathbf{i})|=j}} M_k \leq \frac{(\lfloor \frac{k}{2} \rfloor + 1) N^{\lfloor \frac{k}{2} \rfloor + 1}}{N^{\frac{k}{2}+1}} M_k \\ &= \begin{cases} \frac{(\lfloor \frac{k}{2} \rfloor + 1) M_k}{N^{\frac{1}{2}}} & \text{si } k \text{ es par,} \\ \frac{(\lfloor \frac{k}{2} \rfloor + 1) M_k}{N^{\frac{1}{2}}} & \text{si } k \text{ es impar,} \end{cases} \end{aligned}$$

de donde si k es impar se obtiene el resultado

$$\mathbf{E} \left[\frac{1}{N} \operatorname{tr} \left[(A^N)^k \right] \right] \rightarrow 0, \quad \text{cuando } N \rightarrow \infty.$$

Consideremos ahora el caso k par. Retomemos (8) para expresarlo como

$$\mathbf{E} \left[\frac{1}{N} \operatorname{tr} \left[(A^N)^k \right] \right] = S_1 + S_2$$

donde

$$S_1 = \frac{1}{N^{\frac{k}{2}+1}} \sum_{\substack{G(\mathbf{i}) \\ 1 \leq |\tilde{V}(\mathbf{i})| \leq \frac{k}{2}}} Q(\mathbf{i}) \quad \text{y} \quad S_2 = \frac{1}{N^{\frac{k}{2}+1}} \sum_{\substack{G(\mathbf{i}) \\ |\tilde{V}(\mathbf{i})| = \frac{k}{2}+1}} Q(\mathbf{i}).$$

El término S_1 converge a cero, ya que cuando $N \rightarrow \infty$

$$|S_1| \leq \frac{1}{N^{\frac{k}{2}+1}} \sum_{\substack{G(\mathbf{i}) \\ 1 \leq |\tilde{V}(\mathbf{i})| \leq \frac{k}{2}}} |Q(\mathbf{i})| \leq \frac{k}{2} \frac{N^{\frac{k}{2}} M_k}{N^{\frac{k}{2}+1}} \rightarrow 0.$$

Resta demostrar que S_2 converge a $C_{\frac{k}{2}}$. En este caso $|\tilde{V}(\mathbf{i})| = \frac{k}{2} + 1$, lo cual implica, de acuerdo al lema 4, que $|\tilde{E}(\mathbf{i})| \geq \frac{k}{2}$.

Si $|\tilde{E}(\mathbf{i})| > \frac{k}{2}$ entonces $E(\mathbf{i})$ tiene al menos una diarista que no es paralela a ninguna otra diarista, por el lema 11 $Q(\mathbf{i}) = 0$.

Si $|\tilde{E}(\mathbf{i})| = \frac{k}{2}$, por el lema, 4 $\tilde{G}(\mathbf{i})$ es un árbol, por lo tanto a cada diarista en $E(\mathbf{i})$ le corresponde una diarista paralela en $E(\mathbf{i})$. En este caso $Q(\mathbf{i})$ es de la forma

$$\begin{aligned} Q(\mathbf{i}) &= \mathbf{E} \left(|B_{e_1}|^2 |B_{e_2}|^2 \cdots |B_{e_{\frac{k}{2}}}|^2 \right) = \prod_{e \in \tilde{E}(\mathbf{i})} \mathbf{E} |B_e|^2 \\ &= \prod_{e \in \tilde{E}(\mathbf{i})} \mathbf{E} \left| \sqrt{N} A_e^N \right|^2 = 1. \end{aligned}$$

Cada colección de vértices distintos $\mathbf{i} = (i_1, \dots, i_{\frac{k}{2}+1})$ origina $C_{\frac{k}{2}}$ colecciones de $\frac{k}{2}$ aristas (proposición 6), digamos $\tilde{E}_1(\mathbf{i}), \dots, \tilde{E}_{C_{\frac{k}{2}}}(\mathbf{i})$. A su vez, cada una de estas $C_{\frac{k}{2}}$ colecciones de aristas genera $P_{\frac{k}{2}+1}^N$ árboles, ya que al correr el índice \mathbf{i} sobre $1 \leq i_1, \dots, i_{\frac{k}{2}+1} \leq N$, se generarán

todas las permutaciones de N en $\frac{k}{2} + 1$. Por lo tanto

$$\begin{aligned}
S_2 &= \frac{1}{N^{\frac{k}{2}+1}} \sum_{\substack{G(\mathbf{i}) \\ |\tilde{V}(\mathbf{i})|=\frac{k}{2}+1 \\ |\tilde{E}(\mathbf{i})|=\frac{k}{2}}} Q(\mathbf{i}) \\
&= \frac{1}{N^{\frac{k}{2}+1}} \sum_{j=1}^{C_{\frac{k}{2}}} \sum_{\mathbf{i} \in \tilde{G}_j(\mathbf{i})} Q(\mathbf{i}) \quad \text{con } \tilde{G}_j(\mathbf{i}) = \left(\tilde{V}(\mathbf{i}), \tilde{E}_j(\mathbf{i}) \right) \\
&= \frac{1}{N^{\frac{k}{2}+1}} \sum_{j=1}^{C_{\frac{k}{2}}} \sum_{\mathbf{i} \in \tilde{G}_j(\mathbf{i})} 1 = \frac{C_{\frac{k}{2}}}{N^{\frac{k}{2}+1}} P_{\frac{k}{2}+1}^N,
\end{aligned}$$

de donde $S_2 \rightarrow C_{\frac{k}{2}}$ cuando $N \rightarrow \infty$, en virtud de que $P_{\frac{k}{2}+1}^N / N^{\frac{k}{2}+1} = \frac{N!}{(N-\frac{k}{2}-1)! N^{\frac{k}{2}+1}} = \left(\frac{N}{N}\right) \left(\frac{N-1}{N}\right) \cdots \left(\frac{N-\frac{k}{2}}{N}\right) \rightarrow 1$ cuando $N \rightarrow \infty$. ■

Observación 13 *Si en el teorema 8 las diagonales de las matrices de Wigner son cero, $A_{ii}^N = 0$, $1 \leq i \leq N$, se mantiene la conclusión de la convergencia en esperanza (y también la conclusión del teorema 9 de la convergencia casi segura).*

Esto se puede ver siguiendo los mismos pasos la demostración anterior, ya que los árboles no tienen lazos. Observe que la gráfica del ejemplo 10, la cual no es árbol, tiene a la arista (1, 1) como lazo.

4.4. La convergencia casi segura

Para demostrar la convergencia casi segura del teorema 9 usaremos el lema de Borel-Cantelli, el cual establece que si D_1, D_2, \dots son eventos tales que $P(D_1) + P(D_2) + \dots$ converge entonces $\mathbf{P}\left(\bigcap_{n=1}^{\infty} \bigcup_{N=n}^{\infty} D_N\right) = 0$, ver [10].

Demostración del teorema 9. Más abajo probaremos que la varianza está acotada por

$$var \left(\frac{1}{N} \text{tr} \left[(A^N)^k \right] \right) \leq k \frac{M_k + M_k^2}{N^2}. \quad (9)$$

Definimos para $\varepsilon > 0$,

$$D_{N,\varepsilon} = \left\{ \omega \in \Omega : \left| \frac{1}{N} \text{tr} \left([A^N(\omega)]^k \right) - \mathbf{E} \left(\frac{1}{N} \text{tr} \left[(A^N)^k \right] \right) \right| \geq \varepsilon \right\}.$$

Por la desigualdad de Chebyshev, proposición 1,

$$\mathbf{P}(D_{N,\varepsilon}) \leq \frac{M_k + M_k^2}{\varepsilon^2 N^2},$$

de donde

$$\sum_{N=1}^{\infty} \mathbf{P}(D_{N,\varepsilon}) < \infty$$

y por el lema de Borel-Cantelli

$$\mathbf{P}\left(\bigcap_{n=1}^{\infty} \bigcup_{N=n}^{\infty} D_{N,\varepsilon}\right) = 0.$$

Esto implica que existe un evento Ω_0 con $P(\Omega_0) = 1$ tal que para cada $\omega \in \Omega_0$ existe $N(\omega) \geq 1$ tal que $\left| \frac{1}{N} \text{tr} \left[(A^N(\omega))^k \right] - \mathbf{E} \left[\frac{1}{N} \text{tr} \left[(A^N)^k \right] \right] \right| < \varepsilon$ para $n \geq N(\omega)$. Por lo tanto cuando $N \rightarrow \infty$

$$\left| \frac{1}{N} \text{tr} \left((A^N)^k \right) - \mathbf{E} \left[\frac{1}{N} \text{tr} \left((A^N)^k \right) \right] \right| \xrightarrow{c.s.} 0.$$

Luego del teorema 8 y la desigualdad del triángulo se obtiene $\frac{1}{N} \text{tr} \left((A^N)^k \right) \xrightarrow{c.s.} C_{\frac{k}{2}}$ o 0 de acuerdo a si k es par o impar.

Resta probar (9).

$$\begin{aligned} \text{var} \left(\frac{1}{N} \text{tr} \left[(A^N)^k \right] \right) &= \mathbf{E} \left[\left(\frac{1}{N} \text{tr} \left[(A^N)^k \right] \right)^2 \right] - \left[\frac{1}{N} \mathbf{E} \text{tr} \left[(A^N)^k \right] \right]^2 \\ &= \mathbf{E} \left[\frac{1}{N^{\frac{k}{2}+1}} \sum_{\mathbf{i}} B_{i_1 i_2} B_{i_2 i_3} \cdots B_{i_k i_1} \right]^2 - \left[\frac{1}{N^{\frac{k}{2}+1}} \sum_{\mathbf{i}} \mathbf{E} [B_{i_1 i_2} B_{i_2 i_3} \cdots B_{i_k i_1}] \right]^2 \\ &= \frac{1}{N^{k+2}} \left\{ \sum_{\mathbf{i}, \mathbf{i}'} \mathbf{E} [B_{i_1 i_2} \cdots B_{i_k i_1} B_{i'_1 i'_2} \cdots B_{i'_k i'_1}] \right. \\ &\quad \left. - \mathbf{E} [B_{i_1 i_2} \cdots B_{i_k i_1}] \mathbf{E} [B_{i'_1 i'_2} \cdots B_{i'_k i'_1}] \right\} \\ &= \frac{1}{N^{k+2}} \sum_{\mathbf{i}, \mathbf{i}'} [Q(\mathbf{i}, \mathbf{i}') - Q(\mathbf{i}) Q(\mathbf{i}')] \end{aligned} \quad (10)$$

donde $\mathbf{i} = (i_1, \dots, i_k)$, $\mathbf{i}' = (i'_1, \dots, i'_k)$ y

$$Q(\mathbf{i}, \mathbf{i}') = \mathbf{E} [B_{i_1 i_2} B_{i_2 i_3} \cdots B_{i_k i_1} B_{i'_1 i'_2} B_{i'_2 i'_3} \cdots B_{i'_k i'_1}].$$

De nuevo identificamos a los índices \mathbf{i}, \mathbf{i}' de la suma anterior con la digráfica

$$G(\mathbf{i}, \mathbf{i}') = (V(\mathbf{i}, \mathbf{i}'), E(\mathbf{i}, \mathbf{i}'))$$

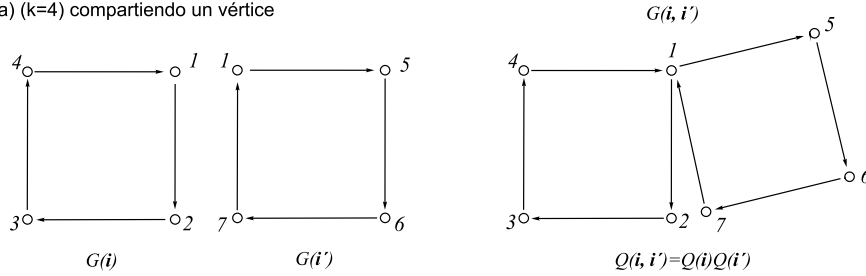
$$V(\mathbf{i}, \mathbf{i}') = \{i_1, \dots, i_k, i'_1, \dots, i'_k\}, \quad 2k \text{ vértices}$$

$$E(\mathbf{i}, \mathbf{i}') = \{(i_1, i_2), \dots, (i_k, i_1), (i'_1, i'_2), \dots, (i'_k, i'_1)\}, \quad 2k \text{ diaristas}$$

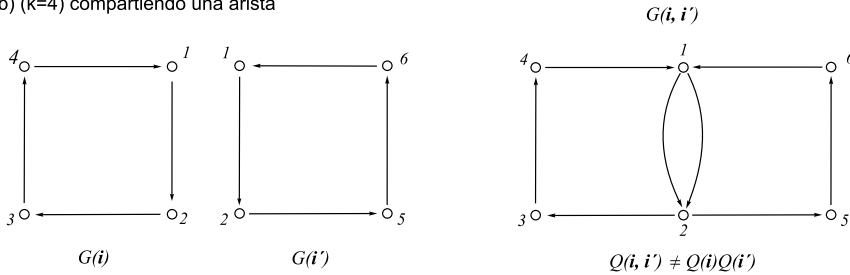
y denotamos por $\tilde{G}(\mathbf{i}, \mathbf{i}') = (\tilde{V}(\mathbf{i}, \mathbf{i}'), \tilde{E}(\mathbf{i}, \mathbf{i}'))$ su correspondiente esqueleto.

Observemos que la digráfica $G(\mathbf{i}, \mathbf{i}')$ no necesariamente es conexa, sin embargo cuando $G(\mathbf{i})$ y $G(\mathbf{i}')$ comparten por lo menos un vértice si es conexa, como lo ilustran las figuras (a) y (b). Además, si los esqueletos $\tilde{G}(\mathbf{i})$ y $\tilde{G}(\mathbf{i}')$ no tienen arista en común, por independencia se sigue que $Q(\mathbf{i}, \mathbf{i}') = Q(\mathbf{i})Q(\mathbf{i}')$.

a) ($k=4$) compartiendo un vértice



b) ($k=4$) compartiendo una arista



De manera similar a la sección anterior analizaremos los términos de la suma (10) agrupándolos de acuerdo al número de vértices de los esqueletos $\tilde{G}(\mathbf{i}, \mathbf{i}')$ correspondientes a estas digráficas $G(\mathbf{i}, \mathbf{i}')$.

Caso $|\tilde{V}(\mathbf{i}, \mathbf{i}')| \geq k + 2$. Si $G(\mathbf{i}, \mathbf{i}')$ es inconexa entonces $\tilde{G}(\mathbf{i})$ y $\tilde{G}(\mathbf{i}')$ no tienen arista en común, por lo tanto $Q(\mathbf{i}, \mathbf{i}') = Q(\mathbf{i})Q(\mathbf{i}')$. Por otro lado, si $G(\mathbf{i}, \mathbf{i}')$ es conexa, entonces $E(\mathbf{i}, \mathbf{i}')$ tiene por lo menos una diarista que no es paralela a ninguna otra diarista, la cual estará en $E(\mathbf{i})$ o $E(\mathbf{i}')$, obteniéndose $Q(\mathbf{i}, \mathbf{i}') = 0$ y $Q(\mathbf{i})Q(\mathbf{i}') = 0$.

Caso $|\tilde{V}(\mathbf{i}, \mathbf{i}')| = k + 1$. Por el lema 4 $|\tilde{E}(\mathbf{i}, \mathbf{i}')| \geq k$. Analicemos por separado estos dos casos:

Si $|\tilde{E}(\mathbf{i}, \mathbf{i}')| = k$ entonces $\tilde{G}(\mathbf{i}, \mathbf{i}')$ es un árbol, luego los árboles $\tilde{G}(\mathbf{i})$ y $\tilde{G}(\mathbf{i}')$ no tienen arista en común, por lo tanto $Q(\mathbf{i}, \mathbf{i}') = Q(\mathbf{i})Q(\mathbf{i}')$. Si $|\tilde{E}(\mathbf{i}, \mathbf{i}')| > k$ entonces $E(\mathbf{i}, \mathbf{i}')$ tiene por lo menos una diarista que no es paralela a ninguna otra diarista, (sígase la demostración del lema 12), la cual estará en $E(\mathbf{i})$ o $E(\mathbf{i}')$, por lo que $Q(\mathbf{i}, \mathbf{i}') = 0$ y $Q(\mathbf{i})Q(\mathbf{i}') = 0$.

Finalmente, en el caso $|\tilde{V}(\mathbf{i}, \mathbf{i}')| \leq k$ la suma (10) se reduce a

$$\begin{aligned} \text{Var} \left(\frac{1}{N} \text{tr} \left[(A^N)^k \right] \right) &= \frac{1}{N^{k+2}} \sum_{\substack{G(\mathbf{i}, \mathbf{i}') \\ 1 \leq |\tilde{V}(\mathbf{i}, \mathbf{i}')| \leq k}}^N [Q(\mathbf{i}, \mathbf{i}') - Q(\mathbf{i})Q(\mathbf{i}')] \\ &\leq \frac{1}{N^{k+2}} \sum_{\substack{G(\mathbf{i}, \mathbf{i}') \\ 1 \leq |\tilde{V}(\mathbf{i}, \mathbf{i}')| \leq k}}^N [|Q(\mathbf{i}, \mathbf{i}')| + |Q(\mathbf{i})Q(\mathbf{i}')|] \\ &\leq \frac{k}{N^{k+2}} N^k (M_k + M_k^2) \\ &= k \frac{M_k + M_k^2}{N^2}. \end{aligned}$$

■

Observación 14 *El teorema 7 es válido para matrices de Wigner tanto reales como complejas, ya que la demostración no distingue el caso real del complejo. El caso real corresponde al de las matrices aleatorias simétricas.*

Si en el teorema 7 se relaja la condición (7b) sobre las varianzas por la condición asintótica

$$\lim_{N \rightarrow \infty} \mathbf{E} \left| \sqrt{N} A_{ij}^N \right|^2 = \lim_{N \rightarrow \infty} \min_{i, j \leq N} \mathbf{E} \left| \sqrt{N} A_{ij}^N \right|^2 = \lim_{N \rightarrow \infty} \max_{i, j \leq N} \mathbf{E} \left| \sqrt{N} A_{ij}^N \right|^2 = 1 \quad (11)$$

se sigue manteniendo la conclusión.

En [4] se considera, en lugar de la condición (7b), la condición

$$\lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{1 \leq i, j \leq N} \left| \mathbf{E} \left[\left| \sqrt{N} A_{ij}^N \right|^2 \right] - 1 \right| = 0 \quad (12)$$

y también se obtiene la conclusión del teorema 7.

Obsérvese que (7b) cumple trivialmente (11) y (12).

5. Notas

Las condiciones (7a), (7b) y (7c) son esencialmente las mismas que considera Wigner [11] para matrices aleatorias simétricas reales, con la diferencia de que considera las entradas con distribución simétrica.

Existen varias versiones del teorema de Wigner dependiendo de las matrices aleatorias que se consideren. Los trabajos que a continuación señalamos emplean diferentes técnicas para demostrarlo. Y aunque trabajan con distintas condiciones, sobre las matrices aleatorias, en general consideran matrices aleatorias simétricas o hermitianas, centradas y con entradas independientes (salvo simetría).

Hiai y Petz [6] trabajan con matrices aleatorias simétricas reales, emplean el enfoque combinatorio, sólo que en lugar de gráficas conexas, utilizan el concepto de particiones que no se cruzan para obtener los números de Catalan.

Nica y Speicher [9] consideran matrices aleatorias con entradas gaussianas complejas y también utilizan particiones que no se cruzan para el conteo. Trabajan el método de momentos con elementos de probabilidad libre y aplican el teorema del límite central libre para obtener el resultado.

Haagerup y Thorbjørnsen [5] también consideran matrices aleatorias con entradas gaussianas complejas. Mediante el enfoque analítico derivan una fórmula para la función generadora de momentos de las matrices aleatorias vía polinomios ortogonales, con la cual deducen el resultado.

Bai y Silverstein [2] consideran matrices aleatorias con entradas complejas a las que les remueven los elementos de la diagonal. Utilizan el enfoque combinatorio de gráficas y la técnica de truncamiento.

Guionnet [4] trabaja con matrices aleatorias hermitianas complejas mediante el enfoque combinatorio de gráficas, de manera similar a como se presenta en este trabajo.

Arnold [1] considera matrices aleatorias simétricas con segundo momento finito en la diagonal y cuarto momento finito fuera de diagonal y demuestra la convergencia en probabilidad. También demuestra la convergencia casi segura, pidiendo cuarto momento finito para los elementos de la diagonal y sexto momento finito fuera de la diagonal. En ambos casos emplea el enfoque combinatorio y la técnica de truncamiento.

Agradecimiento. Al Dr. Víctor Pérez-Abreu por sus comentarios y sugerencias.

Referencias

- [1] Arnold, L. (1967). On the asymptotic distribution of the eigenvalues of random matrices. *Journal of Mathematical Analysis and Applications*, 20, 262-268.
- [2] Bai, Z. y Silverstein, J. W. (2010). *Spectral Analysis of Large Dimensional Random Matrices*. Springer Series in Statistics, 2^a ed.
- [3] Bondy J. A. y Murty U. S. R. (2008) *Graph Theory*. Springer. Graduate Texts in Mathematics 244.
- [4] Guionnet, A. (2009). *Large Random Matrices: Lectures on Macroscopic Asymptotics*. École d'Été de Probabilités de Saint-Flour XXXVI 2006 (Lecture Notes in Mathematics, Ecole d'Été de Probabilités de St. Flour, Springer).
- [5] Haagerup, U. and Thorbjørnsen, S. (2003) Random matrices with complex gaussian entries. *Expositiones Mathematicae*, 21, 293-335.
- [6] Hiai F. y Petz, D. (2000). *The Semicircle Law, Free Random Variables and Entropy*. Mathematics Surveys and Monographs 77, American Mathematical Society.
- [7] Koshy, T. (2009). *Catalan Numbers with applications*. Oxford University Press, Inc.
- [8] Mehta, M. L. (2004). *Random Matrices*. 3^a ed. Elsevier
- [9] Nica, A. y Speicher, R. (2006). *Lectures on the Combinatorics of Free Probability*. London Mathematical Society, Lecture Notes Series 335, Cambridge University Press.
- [10] Rincón, L. (2008). *Curso intermedio de probabilidad*. Editorial Las Prensas de Ciencias, Facultad de Ciencias UNAM, 2^a ed. Versión 2007 en PDF: <http://www.matematicas.unam.mx/lars>
- [11] Wigner, E. P. (1958). On the distributions of the roots of certain symmetric matrices. *Annals of Mathematics*, vol. 67, no. 2, 325-327.