

Comunicaciones del CIMAT

EL TEOREMA DE WIGNER PARA MATRICES ALEATORIAS

J. Armando Domínguez Molina and Alfonso Rocha Arteaga

Comunicación del CIMAT No I-09-08/15-10-2009
(PE/CIMAT)



CIMAT

El Teorema de Wigner para Matrices Aleatorias*

J. Armando Domínguez Molina[†]

jadguez@uas.uasnet.mx

Alfonso Rocha Arteaga

arteaga@uas.uasnet.mx

Escuela de Ciencias Físico-Matemáticas
 Universidad Autónoma de Sinaloa

Resumen

El propósito de este artículo es dar una demostración elemental del Teorema de Wigner, el cual es un resultado central de la teoría de matrices aleatorias, usando herramienta de combinatoria, álgebra lineal y probabilidad de nivel licenciatura.

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 no elementales de este resultado básico de la teoría de matrices aleatorias, sin embargo en este artículo se pretende exponer una demostración detallada usando herramienta elemental de nivel licenciatura.

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 (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 ,

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

*Este trabajo es parte de las exposiciones del Seminario Interinstitucional de Matrices Aleatorias celebrado en la ciudad de Guanajuato, Gto. en julio de 2009, que organizan el Centro de Investigación en Matemáticas, A. C. (CIMAT) y la Universidad Autónoma de Sinaloa (UAS).

[†]Trabajo parcialmente desarrollado por el autor mientras era profesor del Departamento de Matemáticas de la Universidad de Guanajuato.

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 [9] durante la década de los 50's 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 matrices aleatorias que se consideren, [1], [3], [4], [5], [6], [7], [8].

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 $\int x^k F_N(dx)$ y $\int x^k F(dx)$, 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

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

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

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

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

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

Más aún, probaremos la convergencia en esperanza

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

y usaremos este resultado para probar (5). Esta última convergencia la probaremos usando el enfoque combinatorio de gráficas conexas. 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 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 $(A_N)_{ij} = (\bar{A}_N)_{ji}$, con entradas independientes $(A_N)_{ij} \quad 1 \leq i \leq j \leq N$, 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 [9] para matrices aleatorias simétricas reales, salvo por la varianza constante y la distribución simétrica de las entradas.

Este artículo está organizado como sigue. En la Sección 2 se prueba que la distribución del semicírculo está únicamente determinada por sus momentos (3) a través de los números de Catalan. En la Sección 3 se prueba que los números de Catalan (4) cuentan el número de árboles orientados con raíz. Luego se usa este resultado 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 de Teorema de Wigner.

2 Distribución del semicírculo y números de Catalan

Recordemos que cualquier función de distribución que tiene soporte acotado está únicamente determinada por sus momentos, por lo tanto la distribución del semicírculo está únicamente

determinada por sus momentos, sólo resta ver que sus momentos están dados por los números de Catalan.

Proposición 1 *Los momentos pares de la distribución del semicírculo están dados por los números de Catalan (4)*

$$\int_{-2}^2 x^{2k} \frac{1}{2\pi} \sqrt{4-x^2} dx = \frac{1}{k+1} \binom{2k}{k}.$$

Es claro que los momentos impares son cero por la simetría de la distribución.

Demostración. Sea

$$\begin{aligned} m_{2k} &= \frac{1}{2\pi} \int_{-2}^2 x^{2k} \sqrt{4-x^2} dx \\ &= \frac{1}{\pi} \int_0^2 x^{2k} \sqrt{4-x^2} dx \quad \text{simetría} \\ &= \frac{2^{2k+1}}{\pi} \int_0^1 y^{k-\frac{1}{2}} (1-y)^{\frac{1}{2}} dy \quad \text{cambio } x = 2\sqrt{y} \\ &= \frac{2^{2k+1}}{\pi} \frac{\Gamma(k+\frac{1}{2}) \Gamma(\frac{3}{2})}{\Gamma(k+2)} \quad \text{Beta}(\alpha_1, \alpha_2) \quad \alpha_1 = k+1/2 \quad \alpha_2 = 3/2, \end{aligned}$$

luego $\Gamma(k+2) = (k+1)!$ y aplicando iterativamente la propiedad $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$ para $\alpha > 0$,

$$\Gamma(k+1/2) = \frac{(2k-1)(2k-3)(2k-5)\dots 3 \cdot 1}{2^k} \Gamma\left(\frac{3}{2}\right),$$

$$\Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{1}{2}\sqrt{\pi},$$

de donde

$$\begin{aligned} m_{2k} &= \frac{2^{2k+1}}{\pi(k+1)!} \frac{(2k-1)(2k-3)(2k-5)\dots 3 \cdot 1 \sqrt{\pi} \sqrt{\pi}}{2^k} \\ &= \frac{2^k}{(k+1)k!} (2k-1)(2k-3)(2k-5)\dots 1 \\ &= \frac{2^k}{(k+1)k!} \frac{(2k)!}{2^k k!}. \end{aligned}$$

■

3 Árboles y números de Catalan

Ahora veremos que los números de Catalan enumeran los árboles orientados con raíz.

Para esto recordaremos algunos elementos de combinatoria como gráficas, árboles y trayectorias de Dyck. La idea es identificar a los árboles con las trayectorias de Dyck para realizar el conteo sobre estas últimas.

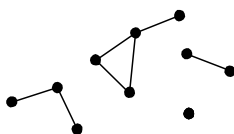
Una *arista* es un par ordenado de puntos llamados *vértices*. Una *gráfica* es un conjunto finito de vértices V y un conjunto de aristas E formadas con dichos vértices

$$V = \{i_1, \dots, i_k\}$$

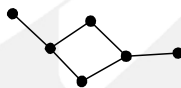
$$E = \{e_i\}_{i \in I} \quad e_i = (i_l, i_m) \quad (l, m) \in \{1, 2, \dots, k\}^2.$$

Una *gráfica conexa* es una gráfica (V, E) tal que cualquier par de vértices v_1, v_2 en V se pueden conectar por una *trayectoria*, esto es, existe una subcolección de aristas $(i_1, i_2), (i_2, i_3), \dots, (i_{n-1}, i_n)$ de E tal que $v_1 = i_1$ y $v_2 = i_n$.

Una gráfica conexa tiene un *bucle* si existe una subcolección de aristas $(i_1, i_2), (i_2, i_3), \dots, (i_{n-1}, i_n)$ de E tal que $i_1 = i_n$. Un *árbol* es una gráfica conexa sin bucles.



gráfica (no conexa)



gráfica conexa con bucles



árbol

El siguiente resultado que se puede ver en [4] y que usaremos en la Sección 4 nos dice que las gráficas conexas no pueden tener más vértices que aristas; salvo el caso de los árboles cuyo número de vértices supera exactamente en uno a su número de aristas. Una *arista no dirigida* es una arista en la que el orden de los vértices no importa, esto es $(i_l, i_m) = (i_m, i_l)$. En cambio, una *arista dirigida* es simplemente una arista, aquí sí importa el orden de los vértices, $(i_l, i_m) \neq (i_m, i_l)$.

Lema 2 Si $G = (V, E)$ es gráfica conexa con aristas no dirigidas entonces

$$|V| \leq |E| + 1.$$

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

Un *árbol con raíz* es un árbol al que se le ha fijado una arista dirigida la cual llamaremos *raíz*. Nos interesa contar el número de árboles con raíz que se pueden formar con k aristas. Una manera de contarlos es identificándolos con $2k$ -tuplas de entradas dicotómicas. Para esto es importante pensar en un árbol como si estuviera en un plano, con la idea de recorrer su contorno partiendo de su raíz y finalizando en ella. Así pues, a cada árbol con raíz de k aristas lo identificamos con la $2k$ -tupla (ξ_1, \dots, ξ_{2k}) construida de la siguiente manera, si en

Proposición 5 *El número de árboles con raíz que se pueden formar con k aristas 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 con raíz que se pueden formar con un número fijo de aristas.

Demostración. Primero contamos el número de trayectorias que inician en $(0, 0)$ y terminan en (m, n) . Cualquiera de ellas que avance u pasos hacia arriba y v hacia abajo termina en el punto $(u + v, u - v)$ de donde

$$u = \frac{m+n}{2} \quad v = \frac{m-n}{2}.$$

El número de trayectorias de $(0, 0)$ a (m, n) que dan $u = \frac{m+n}{2}$ pasos hacia arriba (y por lo tanto $v = \frac{m-n}{2}$ pasos hacia abajo) es el número de m -tuplas de $\{-1, 1\}^m$ que tienen exactamente $\frac{m+n}{2}$ elementos 1,

$$\binom{m}{\frac{m+n}{2}}.$$

En particular el número de trayectorias que van de $(0, 0)$ a $(2k, 0)$ es

$$\binom{2k}{k},$$

sin embargo muchas de estas trayectorias no son de Dyck (las que bajan del eje x).

Afirmación: El número de trayectorias que van de $(0, 0)$ a $(2k, 0)$ que no son de Dyck es igual al número de trayectorias que van de $(0, 0)$ a $(2k, -2)$.

Tal número está dado por

$$\binom{2k}{\frac{2k-2}{2}} = \binom{2k}{k-1}.$$

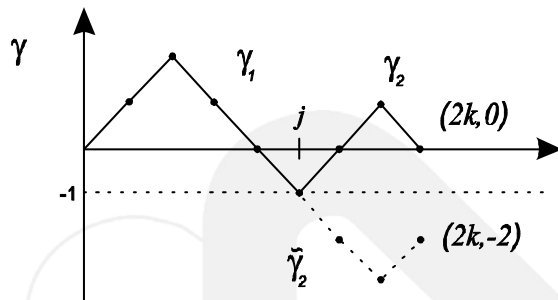
Por lo tanto el número de trayectorias de Dyck de $2k$ pasos es

$$\binom{2k}{k} - \binom{2k}{k-1} = \frac{1}{k+1} \binom{2k}{k}.$$

Resta probar la afirmación. Sea γ una trayectoria que va de $(0, 0)$ a $(2k, 0)$ que no es de Dyck. Sea j el mínimo de $\{1, 2, \dots, 2k-1\}$ tal que $(j, -1) \in \gamma$. Escribamos

$$\begin{aligned} \gamma &= \gamma_1 \vee \gamma_2 && \text{yuxtaposición} \\ \gamma_1 & \text{ trayectoria de } (0, 0) \text{ a } (j, -1) \\ \gamma_2 & \text{ trayectoria de } (j, -1) \text{ a } (2k, 0) \end{aligned}$$

y sea $\tilde{\gamma}_2$ la reflexión de γ_2 sobre la recta horizontal $y = -1$, la cual es una trayectoria de $(j, -1)$ a $(2k, -2)$, como se muestra en la figura.



Definamos

$$\varphi(\gamma) = \gamma_1 \vee \tilde{\gamma}_2$$

la cual es una trayectoria que va de $(0, 0)$ a $(2k, -2)$. Es obvio que φ es uno a uno. Supongamos ahora que β es una trayectoria que va de $(0, 0)$ a $(2k, -2)$. Entonces existe un mínimo $j \in \{1, 2, \dots, 2k - 1\}$ tal que $(j, -1) \in \beta$. Escribamos

$$\begin{aligned} \beta &= \beta_1 \vee \beta_2 && \text{yuxtaposición} \\ \beta_1 &&& \text{trayectoria de } (0, 0) \text{ a } (j, -1) \\ \beta_2 &&& \text{trayectoria de } (j, -1) \text{ a } (2k, -2) \end{aligned}$$

y tomemos la reflexión $\tilde{\beta}_2$ de β_2 sobre $y = -1$, la cual es una trayectoria que va de $(j, -1)$ a $(2k, 0)$.

Sea $\gamma = \beta_1 \vee \tilde{\beta}_2$ yuxtaposición, la cual es una trayectoria que va de $(0, 0)$ a $(2k, 0)$ que no es de Dyck ya que $(j, -1) \in \gamma$. Observemos que

$$\varphi(\gamma) = \beta_1 \vee \beta_2 = \beta,$$

lo cual prueba que φ es sobre. ■

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 $A_{ij}^N = \bar{A}_{ji}^N$.

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

$$\mathbf{E} A_{ij}^N = 0, \tag{7a}$$

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

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

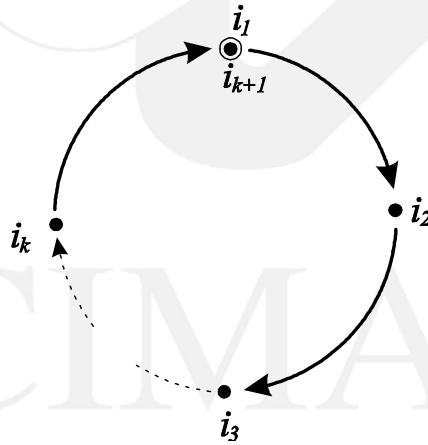
de donde

$$\begin{aligned}
 \mathbf{E} \left[\frac{1}{N} \operatorname{tr} \left[(A^N)^k \right] \right] &= \mathbf{E} \left[\frac{1}{N} \operatorname{tr} \left[\left(N^{-\frac{1}{2}} B^N \right)^k \right] \right] \\
 &= \frac{1}{N^{\frac{k}{2}+1}} \sum_{i_1, \dots, i_k=1}^N \mathbf{E} [B_{i_1 i_2} B_{i_2 i_3} \cdots B_{i_k i_1}] \\
 &= \frac{1}{N^{\frac{k}{2}+1}} \sum_{\mathbf{i}} Q(\mathbf{i}). \tag{8}
 \end{aligned}$$

Vemos de aquí que la estructura 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 gráfica conexa $G(\mathbf{i}) = (V(\mathbf{i}), E(\mathbf{i}))$ definida como

$$\begin{aligned}
 V(\mathbf{i}) &= \{i_1, \dots, i_k\} \text{ vértices.} \\
 E(\mathbf{i}) &= \{(i_1, i_2), \dots, (i_k, i_1)\} \text{ aristas dirigidas.}
 \end{aligned}$$

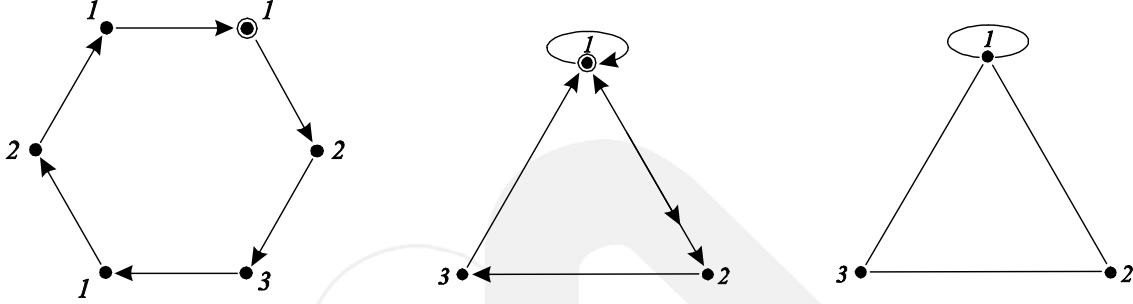
donde $V(\mathbf{i})$ y $E(\mathbf{i})$ admiten elementos repetidos. Esta gráfica conexa se puede representar por la siguiente figura que ilustra la estructura cíclica.



A cada una de estas gráficas conexas $G(\mathbf{i})$ le asociamos su *esqueleto* $\tilde{G}(\mathbf{i})$ definido como la gráfica conexa $\tilde{G}(\mathbf{i}) = (\tilde{V}(\mathbf{i}), \tilde{E}(\mathbf{i}))$ donde $\tilde{V}(\mathbf{i})$ y $\tilde{E}(\mathbf{i})$ no admiten elementos repetidos, esto es $\tilde{V}(\mathbf{i})$ es el conjunto de los distintos vértices de $V(\mathbf{i})$ y $\tilde{E}(\mathbf{i})$ es el conjunto de aristas no dirigidas que se obtiene de quitar las repeticiones de $E(\mathbf{i})$.

Ejemplo 9 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 gráfica $G(\mathbf{i})$ mientras que la tercera es una representación de su esqueleto $\tilde{G}(\mathbf{i})$.

Podemos descomponer la suma (8) clasificando sus términos de acuerdo al número de vértices que tienen los esqueletos de las gráficas

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

Los dos lemas siguientes muestran que esta última suma es cero, ya que sus correspondientes $Q(\mathbf{i})$ no contribuyen a ella. Una arista (i_l, i_{l+1}) de $E(\mathbf{i})$ es una *arista aislada* si para cualquier otra arista (i_m, i_{m+1}) de $E(\mathbf{i})$ se tiene que $(i_l, i_{l+1}) \neq (i_m, i_{m+1})$ y $(i_l, i_{l+1}) \neq (i_{m+1}, i_m)$. En el Ejemplo 9 las aristas $(2, 3)$, $(3, 1)$ y $(1, 1)$ son aristas aisladas de $E(\mathbf{i})$ mientras que $(1, 2)$ y $(2, 1)$ no lo son.

Lema 10 Si $E(\mathbf{i})$ tiene una arista aislada 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 $(i, j) \neq (l, m)$ ó $(i, j) \neq (m, l)$. Luego si (i_l, i_{l+1}) es una arista aislada de $E(\mathbf{i})$,

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

Lema 11 Si $|\tilde{V}(\mathbf{i})| \geq \lfloor \frac{k}{2} \rfloor + 2$ entonces $E(\mathbf{i})$ tiene por lo menos una arista aislada.

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

$$|S_1| \leq \frac{1}{N^{\frac{k}{2}+1}} \sum_{i_1, \dots, i_{\frac{k}{2}}=1}^N |Q(\mathbf{i})| \leq \frac{N^{\frac{k}{2}} M_k}{N^{\frac{k}{2}+1}} \rightarrow 0.$$

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 2, 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 arista aislada y por el Lema 10 $Q(\mathbf{i}) = 0$.

Si $|\tilde{E}(\mathbf{i})| = \frac{k}{2}$ por el Lema 2 $\tilde{G}(\mathbf{i})$ es un árbol, por lo tanto cada arista en $E(\mathbf{i})$ tiene su arista opuesta en $E(\mathbf{i})$. En este caso $Q(\mathbf{i})$ es de la forma

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

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 no dirigidas (Proposición 5), 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

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

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 12 Si en el Teorema 7 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 8 de la convergencia casi segura).

Esto se puede ver siguiendo los mismos pasos la demostración anterior, ya que los árboles no tienen aristas de la forma (i_m, i_{m+1}) con $i_m = i_{m+1}$, en otras palabras los árboles no tienen lazos. Observe que la gráfica del Ejemplo 9, la cual no es árbol, tiene a la arista $(1, 1)$ como lazo.

4.4 La convergencia casi segura

Con el fin de demostrar la convergencia casi segura del Teorema 8 usaremos el Lema de Borel-Cantelli, el cual establece que si C_1, C_2, \dots son eventos tales que $P(C_1) + P(C_2) + \dots$ converge entonces $\mathbf{P}(\bigcap_{n=1}^{\infty} \bigcup_{N=n}^{\infty} C_N) = 0$.

Demostración del Teorema 8. Más abajo probaremos que la varianza está acotada por

$$\text{Var} \left(\frac{1}{N} \text{tr} \left[(A^N)^k \right] \right) \leq \frac{M_k + M_k^2}{N^2}. \quad (10)$$

Si definimos

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

la desigualdad de Chebyshev nos lleva a

$$\mathbf{P}(C_{N,\varepsilon}) \leq \frac{M_k + M_k^2}{\varepsilon^2 N^2}, \quad \text{para } \varepsilon > 0,$$

de donde

$$\sum_{N=1}^{\infty} \mathbf{P}(C_{N,\varepsilon}) \leq \frac{\pi^2 (M_k + M_k^2)}{6\varepsilon^2} < \infty$$

y por el Lema de Borel-Cantelli [2]

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

es decir, para cada $\varepsilon > 0$ la probabilidad de que los eventos $C_{N,\varepsilon}$ ocurran para una infinidad de índices N es cero, por lo tanto cuando $N \rightarrow \infty$

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

Luego del Teorema 7 y la desigualdad del triángulo se obtiene $\frac{1}{N} \text{tr} \left((A^N)^k \right) \xrightarrow{\text{c.s.}} C_{\frac{k}{2}}$ ó 0 de acuerdo a si k es par o impar.

Resta probar (10).

$$\begin{aligned}
 \text{Var} \left(\frac{1}{N} \text{tr} \left[(A^N)^k \right] \right) &= \mathbf{E} \left[\left(\frac{1}{N} \text{tr} \left[(A^N)^k \right] \right)^2 \right] - \left[\frac{1}{N} \mathbf{E} \text{tr} \left[(A^N)^k \right] \right]^2 \\
 &= \mathbf{E} \left[\frac{1}{N^{\frac{k}{2}+1} \sum_{\mathbf{i}} B_{i_1 i_2} B_{i_2 i_3} \cdots B_{i_k i_1}} \right]^2 \\
 &\quad - \left[\frac{1}{N^{\frac{k}{2}+1} \sum_{\mathbf{i}} \mathbf{E} [B_{i_1 i_2} B_{i_2 i_3} \cdots B_{i_k i_1}] \right]^2 \\
 &= \frac{1}{N^{k+2}} \sum_{\mathbf{i}, \mathbf{i}'} \mathbf{E} [B_{i_1 i_2} B_{i_2 i_3} \cdots B_{i_k i_1} B_{i'_1 i'_2} B_{i'_2 i'_3} \cdots B_{i'_k i'_1}] \\
 &\quad - \frac{1}{N^{k+2}} \sum_{\mathbf{i}, \mathbf{i}'} \mathbf{E} [B_{i_1 i_2} B_{i_2 i_3} \cdots B_{i_k i_1}] \mathbf{E} [B_{i'_1 i'_2} B_{i'_2 i'_3} \cdots B_{i'_k i'_1}] \\
 &= \frac{1}{N^{k+2}} \sum_{\mathbf{i}, \mathbf{i}'} [Q(\mathbf{i}, \mathbf{i}') - Q(\mathbf{i}) Q(\mathbf{i}')] \tag{11}
 \end{aligned}$$

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

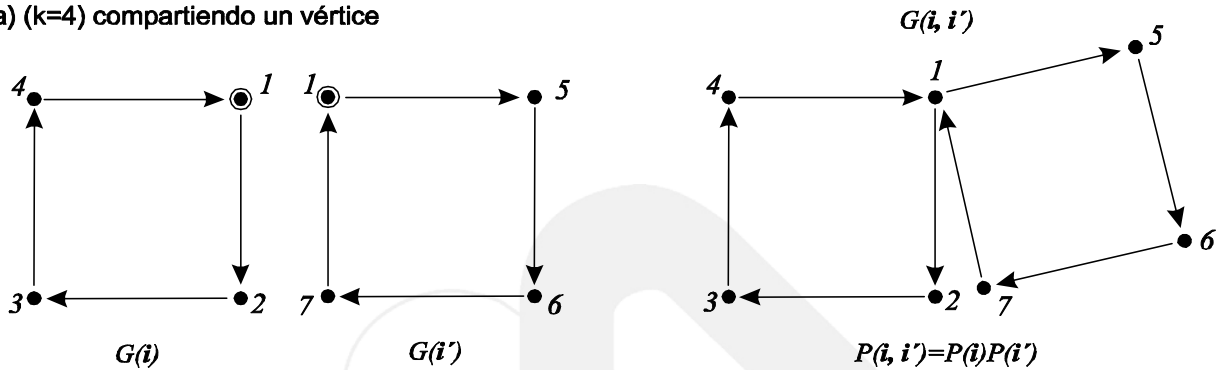
$$Q(\mathbf{i}, \mathbf{i}') = \mathbf{E} [B_{i_1 i_2} B_{i_2 i_3} \cdots B_{i_k i_1} B_{i'_1 i'_2} B_{i'_2 i'_3} \cdots B_{i'_k i'_1}].$$

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

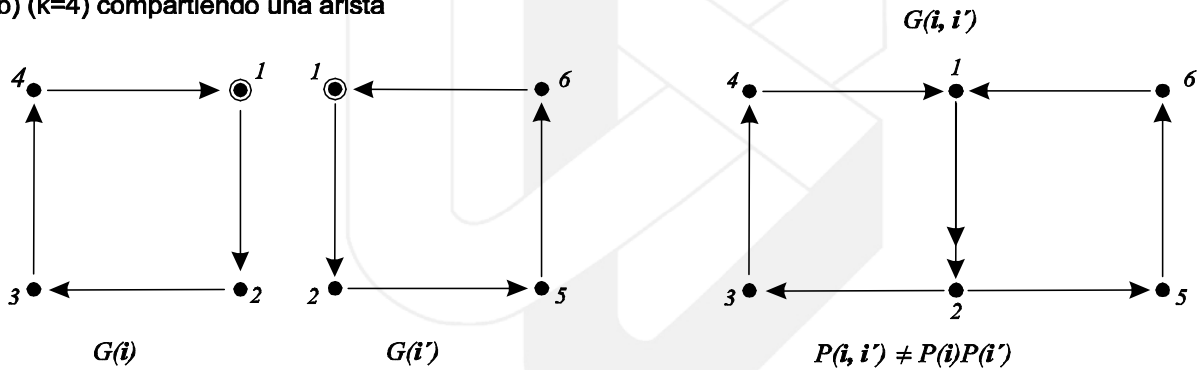
$$\begin{aligned}
 G(\mathbf{i}, \mathbf{i}') &= (V(\mathbf{i}, \mathbf{i}'), E(\mathbf{i}, \mathbf{i}')) \\
 V(\mathbf{i}, \mathbf{i}') &= \{i_1, \dots, i_k, i'_1, \dots, i'_k\}, \quad 2k \text{ vértices} \\
 E(\mathbf{i}, \mathbf{i}') &= \{(i_1, i_2), \dots, (i_k, i_1), (i'_1, i'_2), \dots, (i'_k, i'_1)\}, \quad 2k \text{ aristas.}
 \end{aligned}$$

Observemos que la gráfica $G(\mathbf{i}, \mathbf{i}')$ no necesariamente es conexa, sin embargo cuando $G(\mathbf{i})$ y $G(\mathbf{i}')$ comparten por lo menos un vértice sí es conexa, como lo muestran 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 analizamos los términos de la suma (11) agrupándolos de acuerdo al número de vértices de $\tilde{G}(i, i')$.

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

Caso $|\tilde{V}(i, i')| = k + 1$. Por el Lema 2 $|\tilde{E}(i, i')| \geq k$. Analicemos por separado estos dos casos:

Si $|\tilde{E}(i, i')| = k$ entonces $\tilde{G}(i, i')$ es un árbol, luego los árboles $\tilde{G}(i)$ y $\tilde{G}(i')$ no tienen arista en común, por lo tanto $Q(i, i') = Q(i)Q(i')$.

Si $|\tilde{E}(i, i')| > k$ por lo menos $E(i, i')$ tiene una arista aislada (sígase la demostración del Lema 11), la cual estará en $E(i)$ o $E(i')$, por lo que $Q(i, i') = 0$ y $Q(i)Q(i') = 0$.

Finalmente, en el caso $\left| \tilde{V}(\mathbf{i}, \mathbf{i}') \right| \leq k$ la suma (11) se reduce a

$$\begin{aligned}
 \text{Var} \left(\frac{1}{N} \text{tr} \left[(A^N)^k \right] \right) &= \frac{1}{N^{k+2}} \sum_{\substack{i_1, \dots, i_{k_1}=1 \\ i'_1, \dots, i'_{k_2}=1}}^N [Q(\mathbf{i}, \mathbf{i}') - Q(\mathbf{i})Q(\mathbf{i}')] \quad k_1 + k_2 = k \\
 &\leq \frac{1}{N^{k+2}} \sum_{\substack{i_1, \dots, i_{k_1}=1 \\ i'_1, \dots, i'_{k_2}=1}}^N [|Q(\mathbf{i}, \mathbf{i}')| + |Q(\mathbf{i})Q(\mathbf{i}')|] \\
 &\leq \frac{1}{N^{k+2}} N^k (M_k + M_k^2) \\
 &= \frac{M_k + M_k^2}{N^2}.
 \end{aligned}$$

■

Observación 13 *El Teorema 6 es válido tanto para matrices de Wigner reales como complejas, ya que su demostración no distingue el caso real del complejo. Siendo el caso real, el de las matrices aleatorias simétricas.*

Si en el Teorema 6 se relaja la condición (7b) sobre las varianzas por la condición asintótica

$$\lim_{N \rightarrow \infty} \mathbf{E} \left| \sqrt{N} A_{ij}^N \right|^2 = \lim_{N \rightarrow \infty} \min_{i, j \leq N} \mathbf{E} \left| \sqrt{N} A_{ij}^N \right|^2 = \lim_{N \rightarrow \infty} \max_{i, j \leq N} \mathbf{E} \left| \sqrt{N} A_{ij}^N \right|^2 = 1 \quad (12)$$

se sigue manteniendo la conclusión.

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

$$\lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{1 \leq i, j \leq N} \left| \mathbf{E} \left[\left| \sqrt{N} A_{ij}^N \right|^2 \right] - 1 \right| = 0 \quad (13)$$

y también se obtiene la conclusión del Teorema 6.

Observe que (7b) cumple trivialmente (12) y (13).

5 Notas

Las condiciones (7a), (7b) y (7c) son esencialmente las mismas que considera Wigner [9] para matrices aleatorias simétricas reales, con la diferencia de que considera la varianza constante y 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 consideran otras condiciones, aunque invariablemente trabajan con matrices aleatorias simétricas o hermitianas, centradas y con entradas independientes (salvo simetría).

Por ejemplo, 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 [8] 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 [3] 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 lo presentamos 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.

Agradecimientos. Agradecemos a Víctor Pérez-Abreu por sus comentarios y sugerencias.

Bibliografía

- [1] Arnold, L. (1967). On the asymptotic distribution of the eigenvalues of random matrices. *Journal of Mathematical Analysis and Applications*, 20, 262-268.
- [2] Ash, R. B. y Doléans-Dade, C. A. (2000). *Probability and Measure Theory*. Second Edition, Harcourt-Academic Press.
- [3] Bai, Z. y Silverstein, J. W. (2006). *Spectral Analysis of Large Dimensional Random Matrices*. Science Press, Beijing.
- [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, Vol. 77, American Mathematical Society.

- [7] Mehta, M. L. (2004). *Random Matrices*. Third Edition. Elsevier
- [8] Nica, A. y Speicher, R. (2006). *Lectures on the Combinatorics of Free Probability*. London Mathematical Society, Lecture Notes Series 335, Cambridge University Press.
- [9] Wigner, E. P. (1958). On the distributions of the roots of certain symmetric matrices. *Annals of Mathematics*, Vol. 67, No. 2, 325-327.



CIMAT