Matrices aleatorias
Correlación de series financieras

Dialid Santiago Ramírez

CIMAT

Abril de 2010

Contenido

- El estudio de las propiedades estadísticas de las matrices aleatorias tiene sus orígenes históricos en la física nuclear.

- El estudio de las propiedades estadísticas de las matrices aleatorias tiene sus orígenes históricos en la física nuclear.
- Este problema data de los años cincuenta; los modelos existentes no explicaban de manera adecuada los niveles de energía de un núcleo complejo.

- El estudio de las propiedades estadísticas de las matrices aleatorias tiene sus orígenes históricos en la física nuclear.
- Este problema data de los años cincuenta; los modelos existentes no explicaban de manera adecuada los niveles de energía de un núcleo complejo.
- La Teoría de Matrices Aleatorias desarrollada en ese contexto fue presentada por Wigner, Dyson y Mehta.

- El estudio de las propiedades estadísticas de las matrices aleatorias tiene sus orígenes históricos en la física nuclear.
- Este problema data de los años cincuenta; los modelos existentes no explicaban de manera adecuada los niveles de energía de un núcleo complejo.
- La Teoría de Matrices Aleatorias desarrollada en ese contexto fue presentada por Wigner, Dyson y Mehta.
- Se ha mostrado que estos resultados tienen aplicaciones importantes en el contexto financiero.

- El estudio de las propiedades estadísticas de las matrices aleatorias tiene sus orígenes históricos en la física nuclear.
- Este problema data de los años cincuenta; los modelos existentes no explicaban de manera adecuada los niveles de energía de un núcleo complejo.
- La Teoría de Matrices Aleatorias desarrollada en ese contexto fue presentada por Wigner, Dyson y Mehta.
- Se ha mostrado que estos resultados tienen aplicaciones importantes en el contexto financiero.
- El estudio de matrices de correlación tiene una larga historia en finanzas y es una piedra angular de la teoría de optimización de portafolios de Markowitz.

- El estudio de las propiedades estadísticas de las matrices aleatorias tiene sus orígenes históricos en la física nuclear.
- Este problema data de los años cincuenta; los modelos existentes no explicaban de manera adecuada los niveles de energía de un núcleo complejo.
- La Teoría de Matrices Aleatorias desarrollada en ese contexto fue presentada por Wigner, Dyson y Mehta.
- Se ha mostrado que estos resultados tienen aplicaciones importantes en el contexto financiero.
- El estudio de matrices de correlación tiene una larga historia en finanzas y es una piedra angular de la teoría de optimización de portafolios de Markowitz.
- Es difícil determinar de una manera confiable una matriz de correlación empírica.

- Debido a la gran complejidad de los fenómenos reales uno podría pensar que los modelos matemáticos requieren considerar un gran número de variables para poder hacer una descripción de ellos.

- Debido a la gran complejidad de los fenómenos reales uno podría pensar que los modelos matemáticos requieren considerar un gran número de variables para poder hacer una descripción de ellos.
- Sin embargo, es un fenómeno curioso que el efecto acumulado de muchas variables independientes en un sistema lo vuelve más previsible conforme el número de variables aumenta, en vez de suceder cuando disminuye.

- De hecho, en muchos casos, el comportamiento exacto de cada cada variable se vuelve esencialmente irrelevante (salvo tal vez por algún o algunos parámetros clave); se tiene el mismo comportamiento observado para el sistema en su conjunto independientemente de lo que los componentes individuales están haciendo.

- De hecho, en muchos casos, el comportamiento exacto de cada cada variable se vuelve esencialmente irrelevante (salvo tal vez por algún o algunos parámetros clave); se tiene el mismo comportamiento observado para el sistema en su conjunto independientemente de lo que los componentes individuales están haciendo.
- Este fenómeno se denomina en ocasiones a la **universalidad**, o el **principio de invarianza**.

- Mecánica estadística: El comportamiento de un sistema de N partículas con respecto a los cambios en energía, volumen, etc., parece casi imposible de calcular con precisión cuándo N es grande; puesto que requiere un conocimiento preciso del sistema y sus interacciones.

Universalidad

- Mecánica estadística: El comportamiento de un sistema de N partículas con respecto a los cambios en energía, volumen, etc., parece casi imposible de calcular con precisión cuándo N es grande; puesto que requiere un conocimiento preciso del sistema y sus interacciones.
- Pero en el límite $N \to \infty$, el comportamiento puede ser controlado por sus principales parámetros, tales como la temperatura y la entropía.

- **La ley de los grandes números.** Si $X_1, \cdots, X_n$ son n variables aleatorias independientes e idénticamente distribuidas , entonces la media

$$\frac{X_1 + \cdots + Xn}{n}$$

converge a la medio de cualquiera de las variables X_i, cuando $n \to \infty$.

Principio de invarianza

- En muchos casos, el comportamiento de un una combinación $F(X_1, \cdots, X_n)$ de variables aleatorias i.i.d. $X_1, \cdots, X_n$, para n grande no depende en gran medida de la distribución de X_i, más sí (en algunos casos) de los principales parámetros de esa distribución, tales como la media y la varianza.

Ensamble

Considere el conjunto de matrices reales, simétricas, de dimensión $N \times N$, con entradas determinadas de manera independiente, a partir de una distribución de probabilidad p. Sea A una matriz de este tipo, dada por

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1N} \\ a_{21} & a_{22} & a_{23} & \cdots & a_{2N} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{N1} & a_{N2} & a_{N3} & \cdots & a_{NN} \end{pmatrix} = A^T \tag{1}$$

la densidad de probabilidad de una observación de A es

$$p(A)dA = \prod_{1 \leq i \leq j \leq N} p(a_{ij})da_{ij}. \tag{2}$$

Esto puede interpretarse como la probabilidad de observar una matriz simétrica real donde la probabilidad de que la ij-ésima entrada esté en el intervalo $[a_{ij}, a_{ij} + da_{ij}]$ es $p(a_{ij})da_{ij}$. Más explícitamente

$$\mathrm{P}(A : a_{ij} \in [\alpha_{ij}, \beta_{ij}]) = \prod_{1 \leq i \leq j \leq N} \int_{\alpha_{ij}}^{\beta_{ij}} p(a_{ij})da_{ij}. \tag{3}$$

Ensambles

Una colección de matrices, junto con una densidad de probabilidad, que describe que tan posible es observar una matriz dada, se llama ensamble de matrices (o ensamble de matrices aleatorias).

Definición (Matrices de Wigner reales.)

Para $1 \leq i < j < \infty$ considere $X_{i,j}$ variables aleatorias reales, i.i.d con media 0 y varianza σ^2, y tales que $X_{j,i} = X_{i,j}$. Por otra parte $X_{i,i}$ son v.a. reales i.i.d. con media 0 y varianza 1. Entonces

$$M_n = [X_{i,j}]_{i,j=1}^{n}$$

es una matriz aleatoria simétrica de tamaño $n \times n$.

Definición (Matrices de Wigner reales.)

Para $1 \leq i < j < \infty$ considere $X_{i,j}$ variables aleatorias reales, i.i.d con media 0 y varianza σ^2, y tales que $X_{j,i} = X_{i,j}$. Por otra parte $X_{i,i}$ son v.a. reales i.i.d. con media 0 y varianza 1. Entonces

$$M_n = [X_{i,j}]_{i,j=1}^{n}$$

es una matriz aleatoria simétrica de tamaño $n \times n$.

Definición (Matrices de Wigner complejas.)

Para $1 \leq i < j < \infty$ considere $X_{i,j}$ variables aleatorias (complejas) i.i.d con media 0, varianza σ^2, y tales que $X_{j,i} = \bar{X}_{i,j}$. Por otra parte $X_{i,i}$ son v.a. (reales) i.i.d. con media 0 y varianza 1. Entonces

$$M_n = [X_{i,j}]_{i,j=1}^{n}$$

es una matriz aleatoria hermitiana de tamaño $n \times n$.

Valores propios

En ambos casos las matrices tienen n valores propios reales, a los que denotaremos por

$$\lambda_1(M) \leq \lambda_2(M) \leq \cdots \leq \lambda_n(M).$$

Valores propios

En ambos casos las matrices tienen n valores propios reales, a los que denotaremos por

$$\lambda_1(M) \leq \lambda_2(M) \leq \cdots \leq \lambda_n(M).$$

Valores singulares

En casos más generales, los valores propios pueden ser complejos con lo cual no podrían ser ordenados, sin embargo se pueden definir los valores singulares

$$0 \leq \sigma_1(M) \leq \sigma_2(M) \leq \cdots \leq \sigma_n(M).$$

que no son mas que los eigenvalores de la matriz $(MM^*)^{\frac{1}{2}}$.

Matrices de Wigner

- Los valores propios y singulares esta relacionados con muchas cantidades importantes, tales como:

$$|det(M)| = \prod_{i=1}^{n} |\lambda_i(M)| = \prod_{i=1}^{n} |\sigma_i(M)|$$

- Los valores propios y singulares esta relacionados con muchas cantidades importantes, tales como:

$$|det(M)| = \prod_{i=1}^{n} |\lambda_i(M)| = \prod_{i=1}^{n} |\sigma_i(M)|$$

- o bien

$$tra(M) = \sum_{i=1}^{n} \lambda_i(M).$$

- Los valores propios y singulares esta relacionados con muchas cantidades importantes, tales como:

$$|det(M)| = \prod_{i=1}^{n} |\lambda_i(M)| = \prod_{i=1}^{n} |\sigma_i(M)|$$

- o bien

$$tra(M) = \sum_{i=1}^{n} \lambda_i(M).$$

- Conforme al principio de invarianza, muchos hechos sobre la distribución de los valores propios y singulares de las matrices aleatorias parecen ser universales en el límite $n \to \infty$, que no dependen del modelo preciso.

Conjunto Gaussiano Ortogonal

Conjunto Gaussiano Ortogonal

- Considere matrices reales de Wigner donde $X_{i,j} \sim N(0,1)$ y $X_{i,i} \sim \sqrt{2}N(0,1)$, a este conjunto de matrices se le llama *Ensamble Gausiano Ortogonal.*

Conjunto Gaussiano Ortogonal

- Considere matrices reales de Wigner donde $X_{i,j} \sim N(0,1)$ y $X_{i,i} \sim \sqrt{2}N(0,1)$, a este conjunto de matrices se le llama *Ensamble Gausiano Ortogonal.*
- Una construcción alternativa es considerar $a_{i,j}, i,j \in \mathbb{Z}$ variables aleatorias i.i.d. con distribución normal estándar y $A_n = [a_{i,j}]_{i,j=1}^n$, entonces el GOE es el conjunto de las matrices de la forma

$$M_n = \frac{A_n + A_n^T}{\sqrt{2}}.$$

Algunos de los resultados mas conocidos que muestran la universalidad son

- **Ley del semi-círculo**. A cada matriz aleatoria A se le puedo asociar una medida de probabilidad,**la distribución de probabilidad de valores propios**:

$$\mu_{A,N}(x)dx = \frac{1}{N}\sum_{i=1}^{N}\delta\left(x - \frac{\lambda_i(A)}{2\sqrt{N}}\right). \tag{4}$$

para cada valor propio normalizado $\dfrac{\lambda_i(A)}{2\sqrt{N}}$ se ha puesto una masa de $\frac{1}{N}$; y dado que hay N de estas masas se tiene una distribución de probabilidad.

Algunos de los resultados mas conocidos que muestran la universalidad son

- **Ley del semi-círculo**. A cada matriz aleatoria A se le puedo asociar una medida de probabilidad,**la distribución de probabilidad de valores propios**:

$$\mu_{A,N}(x)dx = \frac{1}{N}\sum_{i=1}^{N}\delta\left(x - \frac{\lambda_i(A)}{2\sqrt{N}}\right). \tag{4}$$

  para cada valor propio normalizado $\dfrac{\lambda_i(A)}{2\sqrt{N}}$ se ha puesto una masa de $\frac{1}{N}$; y dado que hay N de estas masas se tiene una distribución de probabilidad.

- Si $p(x)$ es una distribución de probabilidad, entonces $\int_a^b p(x)dx$ es la probabilidad de observar un valor en $[a, b]$. En este caso $\int_a^b \mu_{A,N}(x)dx$ es el porcentaje de valores propios normalizados dentro del intervalo $[a, b]$, esto es

$$\int_a^b \mu_{A,N}(x)dx = \frac{\#\left\{i : \frac{\lambda_i(A)}{2\sqrt{N}} \in [a, b]\right\}}{N}. \tag{5}$$

- Considere el ensamble de matrices reales simétricas de tamaño $N \times N$ con entradas independientemente elegidas a partir de una densidad de probabilidad $p(x)$ con media 0, varianza 1 y momentos de orden superior finitos.
- Cuando N tiendo a infinito, $\mu_{A,N}(x)$ converge a la densidad del semi-círculo

$$\frac{2}{\pi}\sqrt{1-x^2}.$$

- Este resultado fue establecido por Wigner para el GOE en 1955 y posteriormente generalizado.

Definición

Suponga que $\{X_{ij}, i, j = 1, 2, \cdots, n\}$ es un doble arreglo de variables aleatorias complejas independientes e idénticamente distribuidas con media cero y varianza 1. Sea

$$\mathbf{X} = (\mathbf{X_1}, \cdots, \mathbf{X_n}) \in \mathbb{R}^{\mathbf{p} \times \mathbf{n}}$$

la matriz de varianza se define como

$$\mathbf{S} = \frac{\mathbf{1}}{\mathbf{n}} \mathbf{X} \mathbf{X}^{\mathbf{t}} \in \mathbb{R}^{\mathbf{p} \times \mathbf{p}} \tag{6}$$

y sean

$$\lambda_1, \cdots, \lambda_p$$

denotana sus valores propios. Definimos la medida espectral aleatoria como

$$\mu_n = \frac{1}{p} \sum_{i=1}^{p} \delta_{\lambda_i}.$$

Función de densidad

Sea $\mathbf{S}$ y μ_n como se describieron anteriormente. Si $\dfrac{p}{n} \to y \in (0,1]$ cuando $n \to \infty$ entonces se tiene que

$$\mu_n \to \mu \quad \text{c.s}$$

con

$$\mu(dx) = \frac{1}{2\pi xy}\sqrt{(b-x)(x-a)}1_{([a,b])}dx$$

donde $a(y) = (1-\sqrt{y})^2$ y $b(y) = (1+\sqrt{y})^2$.

- **Ley del menor valor singular.** Dada una matriz aleatoria A de tamaño $N \times N$, con entradas $X_{i,j}$ i.i.d con media 0 varianza 1. El menor valor normalizado singular $\sqrt{n}\sigma_n(A)$ sigue una distribución

$$(1+x)\exp(-x-x^2)dx.$$

- Este resultado fue establecido para el GOE por Edelman en 1991, y en general por Tao-Vu en 2009.

- Dada una matriz aleatoria A real y simétrica, cuyos eigenvalores están dados por

$$\lambda_1(A) \leq \lambda_2(A) \leq \cdots \leq \lambda_N(A). \tag{7}$$

Los espacios entre los eigenvalores (normalizados) están dados por

$$\frac{\lambda_2(A)}{2\sqrt{N}} - \frac{\lambda_1(A)}{2\sqrt{N}}, \frac{\lambda_3(A)}{2\sqrt{N}} - \frac{\lambda_2(A)}{2\sqrt{N}}, \cdots, \frac{\lambda_N(A)}{2\sqrt{N}} - \frac{\lambda_{N-1}(A)}{2\sqrt{N}}. \tag{8}$$

- En 1957 Wigner conjeturo que cuando $N \to \infty$ el espacio entre valores propios normalizados adyacentes, sigue una distribución que viene dada por la siguiente expresión:

$$p_W(s) = \frac{\pi}{2} s \exp\left(-\frac{\pi s^2}{4}\right). \tag{9}$$

Espacios entre valores propios

- Dada una matriz aleatoria A real y simétrica, cuyos eigenvalores están dados por

$$\lambda_1(A) \leq \lambda_2(A) \leq \cdots \leq \lambda_N(A). \tag{10}$$

  Los espacios entre los eigenvalores con respecto al siguiente vecino más cercano (normalizados) están dados por

$$\frac{\lambda_3(A)}{2\sqrt{N}} - \frac{\lambda_1(A)}{2\sqrt{N}}, \frac{\lambda_4(A)}{2\sqrt{N}} - \frac{\lambda_2(A)}{2\sqrt{N}}, \ldots, \frac{\lambda_N(A)}{2\sqrt{N}} - \frac{\lambda_{N-2}(A)}{2\sqrt{N}}. \tag{11}$$

- El espacio entre valores propios con respecto al siguiente vecino más cercano sigue una distribución que viene dada por la siguiente expresión:

$$p_W(s) = \frac{2^{18}}{3^6\pi^3} s^4 \exp\left(-\frac{64}{9\pi} s^2\right). \tag{12}$$

- La mayoría de los instrumentos financieros tienen rendimientos inciertos, por lo que son activos riesgosos.

Conceptos

- La mayoría de los instrumentos financieros tienen rendimientos inciertos, por lo que son activos riesgosos.
- El principal problema que enfrenta un inversionista es la toma de decisiones para la creación de un portafolio.

- La mayoría de los instrumentos financieros tienen rendimientos inciertos, por lo que son activos riesgosos.
- El principal problema que enfrenta un inversionista es la toma de decisiones para la creación de un portafolio.
- Teoría moderna de portafolios.

Conceptos

- La mayoría de los instrumentos financieros tienen rendimientos inciertos, por lo que son activos riesgosos.
- El principal problema que enfrenta un inversionista es la toma de decisiones para la creación de un portafolio.
- Teoría moderna de portafolios.
- Obtener un buen rendimiento minimizando el riesgo.

- Se obtiene al comprar no uno sino varios valores, ya sea del mismo tipo o de diferentes clases.

- Se obtiene al comprar no uno sino varios valores, ya sea del mismo tipo o de diferentes clases.
- Reducir el riesgo.

- Si se tienen N acciones con precio $P_i(t)$ para el activo i en el tiempo t, con, $t = 0, 1, \cdots, T$, el logaritmo de los rendimientos de los activos $S_i(t)$ es:

$$S_i(t) = \ln P_i(t) - \ln P_i(t-1).$$

- Si se tienen N acciones con precio $P_i(t)$ para el activo i en el tiempo t, con, $t = 0, 1, \cdots, T$, el logaritmo de los rendimientos de los activos $S_i(t)$ es:

$$S_i(t) = \ln P_i(t) - \ln P_i(t-1).$$

- Puesto que los niveles de volatilidad (desviación estándar) son diferentes para cada activo, se define el rendimiento normalizado

$$g_i(t) = \frac{S_i(t) - \bar{S_i}(t)}{\sigma_i},$$

- Si se tienen N acciones con precio $P_i(t)$ para el activo i en el tiempo t, con, $t = 0, 1, \cdots, T$, el logaritmo de los rendimientos de los activos $S_i(t)$ es:

$$S_i(t) = \ln P_i(t) - \ln P_i(t-1).$$

- Puesto que los niveles de volatilidad (desviación estándar) son diferentes para cada activo, se define el rendimiento normalizado

$$g_i(t) = \frac{S_i(t) - \bar{S_i}(t)}{\sigma_i},$$

- Donde σ_i es la desviación estándar de S_i y $\bar{S_i}(t)$ la media. Los elementos de la matriz de correlación empírica C son

$$C_{ij} = \frac{1}{T} \sum_{t=1}^{T} g_i(t) g_j(t). \tag{13}$$

- El objetivo es analizar las propiedades estadísticas de una matriz de correlación C utilizando resultados conocidos de matrices aleatorias.
- Se diagonaliza C y se obtienen sus valores propios λ_i, $i = 1, \cdots, N$
- Se compara el comportamiento con las predicciones a partir de las propiedades de universales del GOE.

- La distribución de las diferencias de los valores propios de acuerdo al vecino más cercano $S = \xi_{k+1} - \xi_k$ está dada por la ecuación:

$$P_{\mathrm{GOE}}(S) = \frac{\pi S}{2} \exp\left(-\frac{\pi}{4} S^2\right) \qquad (14)$$

- La distribución de las diferencias de los valores propios de acuerdo al vecino más cercano $S = \xi_{k+1} - \xi_k$ está dada por la ecuación:

$$P_{\mathrm{GOE}}(S) = \frac{\pi S}{2} \exp\left(-\frac{\pi}{4} S^2\right) \tag{14}$$

- La distribución de las diferencias de los valores propios de acuerdo al siguiente vecino más cercano $u = \xi_{k+2} - \xi_k$ es

$$P_{\mathrm{GOE}}(u) = \frac{2^{18}}{3^6 \pi^3} u^4 \exp\left(-\frac{64}{9\pi} u^2\right).$$

1. **Correlaciones entre valores propios de rango largo.** Para probar correlaciones entre pares de valores propios en rangos largos, se usará el estadístico Σ^2 conocido como *varianza número*, que está definido como la varianza del número de valores propios desplegados en intervalos de longitud l alrededor de cada ξ_i, es decir:

$$\Sigma^2 = \langle [n(\xi, l) - l^2] \rangle_\xi$$

donde $n(\xi, l)$ es el número de valores propios desplegados en el intervalo $[\xi - \frac{1}{2}, \xi + \frac{1}{2}]$ y $\langle \rangle_\xi$ es el promedio sobre todos los ξ.

Predicciones de la teoría de matrices aleatorias

1. **Correlaciones entre valores propios de rango largo.** Para probar correlaciones entre pares de valores propios en rangos largos, se usará el estadístico Σ^2 conocido como *varianza número*, que está definido como la varianza del número de valores propios desplegados en intervalos de longitud l alrededor de cada ξ_i, es decir:

   $$\Sigma^2 = \langle [n(\xi, l) - l^2] \rangle_\xi$$

   donde $n(\xi, l)$ es el número de valores propios desplegados en el intervalo $[\xi - \frac{1}{2}, \xi + \frac{1}{2}]$ y $\langle\rangle_\xi$ es el promedio sobre todos los ξ.
2. Si los valores propios no estan correlacionados $\Sigma^2 \approx l$.

1. **Correlaciones entre valores propios de rango largo.** Para probar correlaciones entre pares de valores propios en rangos largos, se usará el estadístico Σ^2 conocido como *varianza número*, que está definido como la varianza del número de valores propios desplegados en intervalos de longitud l alrededor de cada ξ_i, es decir:

$$\Sigma^2 = \langle [n(\xi, l) - l^2] \rangle_\xi$$

donde $n(\xi, l)$ es el número de valores propios desplegados en el intervalo $[\xi - \frac{1}{2}, \xi + \frac{1}{2}]$ y $\langle \rangle_\xi$ es el promedio sobre todos los ξ.

2. Si los valores propios no estan correlacionados $\Sigma^2 \approx l$.
3. De lo contrario Σ^2 será constante.

Predicciones sobre el comportamiento de los vectores propios

- También se propone analizar el comportamiento de los vectores propios de la matriz.

Predicciones sobre el comportamiento de los vectores propios

- También se propone analizar el comportamiento de los vectores propios de la matriz.
- Los componentes $\{u_l^k; l = 1, \cdots, N\}$ del vector propio u^k de una matriz aleatoriade correlación C siguen una distribución normal estándar.

- También se propone analizar el comportamiento de los vectores propios de la matriz.
- Los componentes $\{u_l^k; l = 1, \cdots, N\}$ del vector propio u^k de una matriz aleatoriade correlación C siguen una distribución normal estándar.
- Para cuantificar el npumero de componentes que participan *significativamente* en cada vector propio, se utiliza el cociente inverso de participación, que esta dado por

$$I^k = \sum_{i=1}^{N} (u_l^k)^4$$

donde N es el número de series de tiempo (empresas) y por tanto el número de componentes.

- También se propone analizar el comportamiento de los vectores propios de la matriz.
- Los componentes $\{u_l^k; l = 1, \cdots, N\}$ del vector propio u^k de una matriz aleatoriade correlación C siguen una distribución normal estándar.
- Para cuantificar el npumero de componentes que participan *significativamente* en cada vector propio, se utiliza el cociente inverso de participación, que esta dado por

$$I^k = \sum_{i=1}^{N} (u_l^k)^4$$

donde N es el número de series de tiempo (empresas) y por tanto el número de componentes.

- Si todas las componentes son idénticas y $u_l^k = \frac{1}{\sqrt{N}}$ se tiene que $I^k = \frac{1}{N}$.

Predicciones sobre el comportamiento de los vectores propios

- También se propone analizar el comportamiento de los vectores propios de la matriz.
- Los componentes $\{u_l^k; l = 1, \cdots, N\}$ del vector propio u^k de una matriz aleatoriade correlación C siguen una distribución normal estándar.
- Para cuantificar el npumero de componentes que participan *significativamente* en cada vector propio, se utiliza el cociente inverso de participación, que esta dado por

$$I^k = \sum_{i=1}^{N} (u_l^k)^4$$

  donde N es el número de series de tiempo (empresas) y por tanto el número de componentes.
- Si todas las componentes son idénticas y $u_l^k = \frac{1}{\sqrt{N}}$ se tiene que $I^k = \frac{1}{N}$.
- Si sólo una componente es distinta de cero $u_l^k = 1$, $I^k = 1$.

- También se propone analizar el comportamiento de los vectores propios de la matriz.
- Los componentes $\{u_l^k; l = 1, \cdots, N\}$ del vector propio u^k de una matriz aleatoriade correlación C siguen una distribución normal estándar.
- Para cuantificar el npumero de componentes que participan *significativamente* en cada vector propio, se utiliza el cociente inverso de participación, que esta dado por

$$I^k = \sum_{i=1}^{N} (u_l^k)^4$$

donde N es el número de series de tiempo (empresas) y por tanto el número de componentes.

- Si todas las componentes son idénticas y $u_l^k = \frac{1}{\sqrt{N}}$ se tiene que $I^k = \frac{1}{N}$.
- Si sólo una componente es distinta de cero $u_l^k = 1$, $I^k = 1$.
- De modo que el CIP es el recíproco del número de componentes del vector que contribuyen significativamente.

Caso Particular BMV

- Las series de tiempo que conforman la base de datos para este estudio están formadas por los precios de cierre diario de 65 empresas que cotizan en la BMV.

- Las series de tiempo que conforman la base de datos para este estudio están formadas por los precios de cierre diario de 65 empresas que cotizan en la BMV.
- Durante un periodo de 8 años.

Caso Particular BMV

- Las series de tiempo que conforman la base de datos para este estudio están formadas por los precios de cierre diario de 65 empresas que cotizan en la BMV.
- Durante un periodo de 8 años.
- Para la elección de las empresas y la longitud de la serie se tuvo en cuenta la bursatilidad y capitalización.

- Las series de tiempo que conforman la base de datos para este estudio están formadas por los precios de cierre diario de 65 empresas que cotizan en la BMV.
- Durante un periodo de 8 años.
- Para la elección de las empresas y la longitud de la serie se tuvo en cuenta la bursatilidad y capitalización.
- La longitud final de las series es de 1598.

- Las series de tiempo que conforman la base de datos para este estudio están formadas por los precios de cierre diario de 65 empresas que cotizan en la BMV.
- Durante un periodo de 8 años.
- Para la elección de las empresas y la longitud de la serie se tuvo en cuenta la bursatilidad y capitalización.
- La longitud final de las series es de 1598.
- Dentro de las 65 empresas seleccionadas para el estudio se encuentran representados todos los sectores económicos, las empresas elegidas tienen la mayor bursatilidad de cada sector y juntas representan más del 85 % de participación en el IPC y el 100 % del índice México (INMEX).

- Las series de tiempo que conforman la base de datos para este estudio están formadas por los precios de cierre diario de 65 empresas que cotizan en la BMV.
- Durante un periodo de 8 años.
- Para la elección de las empresas y la longitud de la serie se tuvo en cuenta la bursatilidad y capitalización.
- La longitud final de las series es de 1598.
- Dentro de las 65 empresas seleccionadas para el estudio se encuentran representados todos los sectores económicos, las empresas elegidas tienen la mayor bursatilidad de cada sector y juntas representan más del 85 % de participación en el IPC y el 100 % del índice México (INMEX).
- Se comparan las distribuciones empíricas de los valores propios y las estadísticas de los vectores propios de la matriz C construida a partir de estos datos con sus predicciones teóricas, asumiendo que la matriz de correlación es puramente aleatoria

Conclusiones

- Se ha encontradó que la mayoría de los valores propios en el espectro de la matriz de correlación C coinciden notablemente bien con las predicciones universales de la Teoría de Matrices Aleatorias.

Conclusiones

- Se ha encontradó que la mayoría de los valores propios en el espectro de la matriz de correlación C coinciden notablemente bien con las predicciones universales de la Teoría de Matrices Aleatorias.
- En particular, se ha encontrado que la matriz C satisface las propiedades universales del conjunto gaussiano ortogonal de matrices simétricas aleatorias.

Conclusiones

- Se ha encontradó que la mayoría de los valores propios en el espectro de la matriz de correlación C coinciden notablemente bien con las predicciones universales de la Teoría de Matrices Aleatorias.
- En particular, se ha encontrado que la matriz C satisface las propiedades universales del conjunto gaussiano ortogonal de matrices simétricas aleatorias.
- El cociente inverso de participación soporta la idea de que algunas acciones dominan el mercado y más especfícamente nos dice que el vector propio u_{65} contiene aproximadamente $\frac{1}{I_{65}} = 40$ participantes significativos, que son precisamente las acciones con mayor capitalización en el mercado.

Conclusiones

- Se ha encontradó que la mayoría de los valores propios en el espectro de la matriz de correlación C coinciden notablemente bien con las predicciones universales de la Teoría de Matrices Aleatorias.
- En particular, se ha encontrado que la matriz C satisface las propiedades universales del conjunto gaussiano ortogonal de matrices simétricas aleatorias.
- El cociente inverso de participación soporta la idea de que algunas acciones dominan el mercado y más especfícamente nos dice que el vector propio u_{65} contiene aproximadamente $\frac{1}{I_{65}} = 40$ participantes significativos, que son precisamente las acciones con mayor capitalización en el mercado.
- Estos resultados son cuestionables pues también se ha visto que para series de tiempo de menor longitud, las matrices de correlación no satisfacen las propiedades universales.

Referencias

- Laloux, L. Cizeau, P. Bouchaud, (2002), *Random Matrix Theory and Financial Correlations.*
- Medina Herrera (2007) *Teoría de matrices aleatorias y correlación de series financieras: El caso de la bolsa mexicana de valores.* (Journal of Management,Finance and Economics.)
- Vasiliky Plerou,Bernd Rosenow, (1999) *Universal and nonuniversal properties of cross correlations in financial time series.*, Physical review letters.