



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS FÍSICAS
INSTITUTO DE CIENCIAS FÍSICAS

**TÉCNICAS NUEVAS EN EL ANÁLISIS DE MERCADOS DE
VALORES**

TESIS
QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS (FÍSICA)

PRESENTA:
JOSÉ FRANCISCO MORALES HERNÁNDEZ

TUTOR PRINCIPAL:
DR. FRANCOIS ALAIN LEYVRAZ WALTZ
INSTITUTO DE CIENCIAS FÍSICAS

MIEMBROS DEL COMITÉ TUTOR:
DR. THOMAS HENRY SELIGMAN SCHURCH
INSTITUTO DE CIENCIAS FÍSICAS
DR. CARLOS FRANCISCO PINEDA ZORRILLA
INSTITUTO DE FÍSICA

MÉXICO D.F., ENERO 2016



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Resumen

La teoría de matrices aleatorias (RMT) se ha aplicado en distintas áreas de conocimiento desde que se hizo popular con los trabajos de Wigner, una de esas áreas, y en la que nos enfocaremos, es la del estudio de series de tiempo de rendimientos. Para estudiar estas series de tiempo se analiza la matriz de correlación, donde el tema apropiado en RMT son los ensembles de Wishart.

En este trabajo se aplica un método nuevo para analizar series de tiempo cortas con el que se obtienen más valores propios. Este método consiste en crear un ensemble de matrices de correlación a partir de un número fijo de series de tiempo y con ésto se puede hacer un análisis de las propiedades estadísticas del espectro de valores propios de dichas matrices. Después se compara con el caso no correlacionado, que corresponde al ensemble de Wishart.

Además de la correlación de Pearson, se estudió la correlación parcial entre las series de tiempo, y se encontró que las fluctuaciones de los espectros de valores propios obtenidos con estos tipos de correlaciones se comportan cualitativamente de la misma forma, pero las distribuciones espectrales son diferentes. Se observó que la estructura de la matriz de correlación juega un papel en los momentos de crisis, y no solo importa el nivel de correlaciones global del mercado.

Índice general

Resumen	I
1. Introducción	1
2. Conceptos fundamentales	3
2.1. Conceptos financieros	3
2.1.1. ¿Qué es un sistema financiero?	3
2.1.2. Mercado primario vs Mercado secundario	4
2.1.3. Índice financiero	5
2.2. Teoría de Matrices Aleatorias (RMT)	6
2.2.1. Ensembles gaussianos	6
2.2.2. Ensemble de Wishart	8
2.2.3. Ensemble de Wishart correlacionado	9
2.2.4. Distribución de espaciamiento a primer vecino	10
2.2.5. Varianza de número	11
3. Descripción del modelo	13
3.1. Modelo	15
3.2. Correlaciones parciales	17
4. Resultados	19
4.1. Distribución de valores propios	20
4.1.1. Bulk	21
4.1.2. Valor propio más grande	24
4.1.3. Efecto de la estructura de la matriz de correlación	27
4.2. Correlación promedio vs Valor propio más grande	32
4.2.1. $\langle C_{ij} \rangle$ vs λ_1	32
4.2.2. $\langle C_{ij} \rangle$ vs $\langle \lambda_1 \rangle_{R,S}$	34
4.2.3. Relación entre λ_1 y $\langle \lambda_1 \rangle_{R,S}$	34
4.3. Distribución de espaciamiento a primer vecino	35
4.4. Varianza de número	36

IV

ÍNDICE GENERAL

5. Conclusiones	41
A. Desdoblamiento	43
Bibliografía	45

Capítulo 1

Introducción

A lo largo de los años se han usado una gran variedad de métodos de análisis de series de tiempo para estudiar el comportamiento de acciones financieras, con el propósito de comprender la estructura y dinámica del mercado y detectar hechos estilizados¹ en el mismo [1]. En la caracterización del perfil de correlaciones, estos métodos generalmente trabajan con el coeficiente de correlación de Pearson, para estudiar la relación entre acciones.

Propiedades emergentes, tales como la presencia de estados del mercado, se han detectado mediante la agrupación de matrices de correlación para distintos tiempos [2]. Otro método, además del estudio de los elementos de correlación, se ha enfocado en la aplicación [3–7] de Teoría de Matrices Aleatorias (RMT) [8, 9], principalmente en el estudio de la distribución espectral de matrices grandes de correlación. Generalmente el problema con estos análisis es que, debido a que se estudian series de tiempo muy largas, no se pueden caracterizar las propiedades de las distribuciones espectrales en periodos de tiempo corto, además de que las series de tiempo financieras no son estacionarias.

En el trabajo de Münnix et al. [2], a pesar de que se estudian series de tiempo cortas, no se puede hacer un análisis espectral, puesto que mientras más cortas las series de tiempo, menos valores propios diferentes de cero hay. Motivados por lo anterior, en este trabajo se implementa un nuevo método que nos permitirá estudiar series de tiempo cortas, con la ventaja de tener suficientes valores propios que nos permiten hacer un análisis estadístico. Este método consiste en crear un *ensemble* de matrices de correlación $\{C_\alpha\}$, a partir del conjunto de series de tiempo que se quieren caracterizar. Las series de tiempo que analizamos son las series de rendimientos para un conjunto de acciones del índice S&P 500, que, al igual que en [2], fueron seleccionadas

¹ Propiedades estadísticas que son comunes para un amplio rango de instrumentos financieros, mercados y periodos de tiempo.

por haber existido a lo largo del tiempo bajo consideraciones económicas similares.

En el capítulo 2 damos algunos conceptos necesarios para entender los datos con los que se trabaja, es decir, conceptos relacionados con los mercados financieros. Además, se introducen algunos modelos y distribuciones en RMT, principalmente modelos de ensembles de matrices de correlación, como el ensemble de Wishart [10, 11] y Wishart correlacionado [12–14], así como la distribución de espaciamiento a primer vecino $p(s)$ y la varianza de número $\Sigma^2(L)$.

En el capítulo 3 describimos nuestro modelo, que consiste en crear un ensemble de matrices de correlación $\{\mathcal{C}_\alpha\}$ a partir de un conjunto de N series de tiempo de rendimientos. Para crear un miembro del ensemble se eligen $n < N$ series de tiempo de las N originales y consideramos dos formas de elección: aleatoriamente y por sectores, éste último es conservando las proporciones de acciones que corresponden a los diferentes sectores industriales. Además de estudiar la correlación de Pearson, también se analiza la correlación parcial [15] entre las acciones dado el valor promedio de rendimientos del conjunto de acciones, es decir, se remueve el efecto del mercado. Al final de este capítulo se explica cómo se define la correlación parcial entre dos series de tiempo dado una tercera.

En el capítulo 4 se describen los resultados obtenidos con ambos métodos de elección durante la construcción del ensemble, tanto con correlación ordinaria como con correlación parcial. Un resultado importante es que las distribuciones $p(s)$ y $\Sigma^2(L)$ se comportan de forma similar para correlación ordinaria y parcial, a pesar de que el promedio del valor propio más grande $\langle \lambda_1 \rangle$ difiere substancialmente. Además, la forma de elección en la construcción del ensemble no juega un papel importante en el método.

Finalmente en el capítulo 5 se presentan las conclusiones del trabajo desarrollado.

Capítulo 2

Conceptos fundamentales

2.1. Conceptos financieros

2.1.1. ¿Qué es un sistema financiero?

Un sistema financiero consiste de instituciones financieras¹ y mercados que interactúan, generalmente de forma compleja. Las instituciones son los agentes principales en los mercados financieros debido a que desempeñan la función de intermediarios.

Un mercado financiero es un mercado donde instrumentos financieros son intercambiados o comerciados por agentes financieros. Estos son bienes intangibles, de los cuales se esperan que provean beneficios futuros, por ejemplo, una ganancia monetaria. Un ejemplo de instrumento financiero son las acciones, las cuales son emitidas por compañías como una forma de obtener fondos y dan a los accionistas derechos sobre los bienes y ganancias de la compañía.

Cualquier transacción relacionada con un instrumento financiero incluye al menos dos agentes: compradores y vendedores. Estos proveen, como principales características, las siguientes tres funciones:

1. Descubrimiento de precio
2. Liquidez
3. Reducción de costos de transacción

La función de descubrimiento de precio se refiere a que las transacciones entre los compradores y vendedores de instrumentos financieros en un

¹ Firmas que proveen servicios financieros a sus consumidores. Por ejemplo, invirtiendo sus fondos.

mercado financiero determinan el precio del bien intercambiado. Esto es un principio económico básico llamado oferta y demanda. Por ejemplo, si los agentes financieros piensan que el precio de una acción es más bajo del que debería, ésto incrementa el interés en el bien, y por lo tanto habrá más compradores que vendedores. Ésto hace que el precio del bien suba.

La función de liquidez provee a los inversionistas la oportunidad de vender instrumentos financieros, ya que se refiere a la capacidad de vender un bien en el mercado en cualquier momento. Sin liquidez, un inversor sería forzado a retener el instrumento financiero hasta que una oportunidad se presente. Todos los mercados financieros tienen algún grado de liquidez, sin embargo, diferentes mercados están caracterizados por ésta.

2.1.2. Mercado primario vs Mercado secundario

Cuando se crean instrumentos financieros y se venden por primera vez, se hace en el llamado *mercado primario*. Por ejemplo, cuando se crean acciones con el objetivo de venderlas, estas se venden por primera vez en un mercado primario. A partir de la primera venta, las ventas subsecuentes se realizan en el *mercado secundario*. Los vendedores y compradores que determinan el precio de los instrumentos interactúan en este tipo de mercado. Las instituciones emisoras no obtienen nuevos fondos cuando los instrumentos se comercian en el mercado secundario; los fondos únicamente son obtenidos en el mercado primario.

Las características de los mercados financieros que se mencionaron en la sección anterior corresponden a características de los mercados secundarios, y generalmente cuando se habla de mercados financieros se piensa en este tipo de mercados.

Los instrumentos financieros pueden ser intercambiados en dos formas. La primera es una bolsa (*exchange*), las bolsas son mercados centralizados y regulados, donde los instrumentos son comerciados de manera estandarizada y transparente. Ejemplos de estas bolsas son: New York Stock Exchange (NYSE), Nasdaq, y la Bolsa Mexicana de Valores (BMV). Generalmente las compañías grandes eligen éste tipo de intercambio para *cotizar* y comerciar sus instrumentos. Sin embargo, muchas compañías no cumplen con los requisitos para cotizar en éste tipo de bolsas o no quieren pagar los costos. Estas empresas cotizan en mercados *extra-bursátiles* (*Over-The-Counter OTC*), el segundo método para intercambiar instrumentos financieros. Los mercados extra-bursátiles son redes descentralizadas de *dealers*, en donde los *brokers* y *dealers* interactúan directamente entre ellos ya sea por redes computacionales o por teléfono. Ésta forma de cotizar es menos transparente ya que, a diferencia de la bolsa, los precios de postura generalmente no son publicados;

a menos que se trate de acciones que también están registradas en la bolsa. Por esta razón, generalmente se analizan los instrumentos financieros que cotizan en la bolsa, como en el caso de este trabajo.

2.1.3. Índice financiero

Un índice financiero es una medida del desempeño de un portafolio o conjunto de instrumentos financieros en un mercado en particular, o de cierta parte del mercado. Típicamente, un índice es un promedio ponderado sobre un conjunto de instrumentos que da información de los cambios en los precios en el portafolio. Además, cada índice es calculado y ponderado de forma diferente.

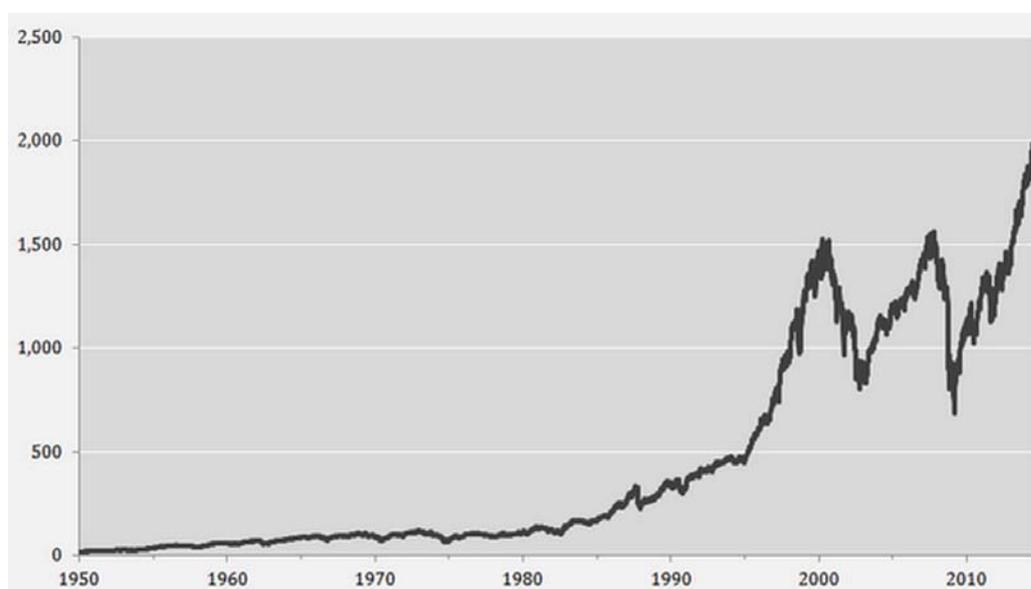


Figura 2.1: Índice S&P 500 desde 1950.

Hay muchos tipos de índices, dependiendo del tipo del mercado o del sector. Uno de los índices financieros más populares es el Standard & Poor's 500 (S&P500) y contiene 500 acciones consideradas como las más representativas de las bolsas NYSE y Nasdaq. Este índice se usa comúnmente como referencia del mercado estadounidense. Ejemplos de otros índices son: Nikkei 225 y FTSE 100, que son índices basados en acciones de Japón y Reino Unido, respectivamente. En México, está el Índice de Precios y Cotizaciones (IPC), que mide el desempeño de 35 acciones que cotizan en la Bolsa Mexicana de Valores (BMV).

2.2. Teoría de Matrices Aleatorias (RMT)

Originalmente la teoría de matrices aleatorias (RMT²) fue aplicada por Wigner [16–18] para tratar con la estadística de eigenvalores y eigenfunciones de sistemas cuánticos complejos de muchos cuerpos. Por ejemplo, ha sido aplicada exitosamente a las propiedades de fluctuaciones espectrales del núcleo atómico y se han estudiado las fluctuaciones estadísticas de los procesos de dispersión en tales sistemas [8, 9]. Se considera que el campo de estudio de RMT comenzó con estos trabajos debido a su influencia, pero su origen se remonta a los trabajos de Wishart [10] y de James [19].

El modelo introducido por Wigner difiere en forma fundamental de la aplicación estándar de conceptos estadísticos en la física. En la mecánica estadística, uno considera un ensemble de sistemas físicos idénticos, todos gobernados por el mismo hamiltoniano pero con condiciones iniciales distintas, y calcula funciones termodinámicas promediando sobre este ensemble. Wigner procedió diferente: él consideró ensembles de sistemas dinámicos gobernados por hamiltonianos *diferentes* con alguna propiedad de simetría en común. El modelo enfoca su atención en las propiedades que son comunes en casi todos los miembros del ensemble y que están determinadas por las simetrías fundamentales sobresalientes.

RMT ha experimentado un desarrollo rápido e inesperado en el último cuarto de siglo. Ha sido aplicada exitosamente a una gran variedad de problemas físicos, y ha llegado a ser una herramienta importante en el estudio de sistemas que son aparentemente bastante diferentes de sistemas de muchos cuerpos, como sistemas moleculares. Por ejemplo: sistemas cuánticos desordenados, sistemas cuánticos caóticos [9] y sistemas financieros. En las figuras 2.2 y 2.3 se muestran ejemplos, de una distribución que caracteriza las fluctuaciones espectrales, para dos sistemas diferentes.

En la siguiente sección se dará una breve descripción del ensemble gaussiano, seguido del ensemble de Wishart y después del de Wishart correlacionado.

2.2.1. Ensembles gaussianos

En RMT, se empieza con un ensemble de hamiltonianos que es definido por un ensemble de matrices aleatorias, donde cada miembro describe una dinámica diferente. En la práctica se usan matrices hamiltonianas de dimensión finita D , aunque para resultados analíticos a menudo se toma el límite $D \rightarrow \infty$.

² Random Matrix Theory.

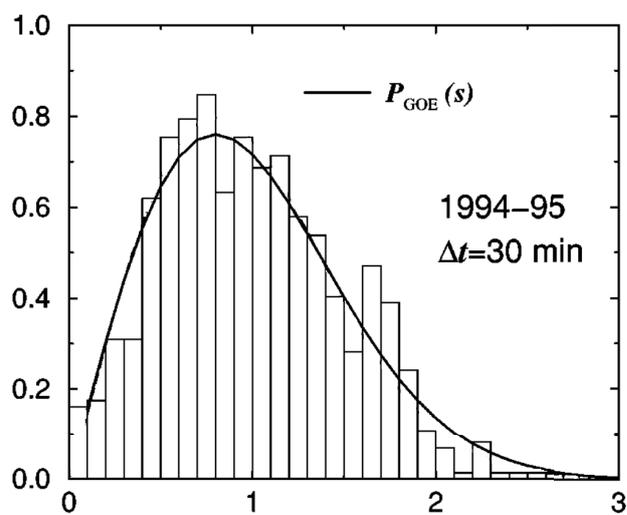


Figura 2.2: Distribución de espaciamiento a primer vecino para una matriz de correlación construida con datos entre 1994 y 1995 para 1000 acciones. En pasos de 30 min. Tomada de [3].

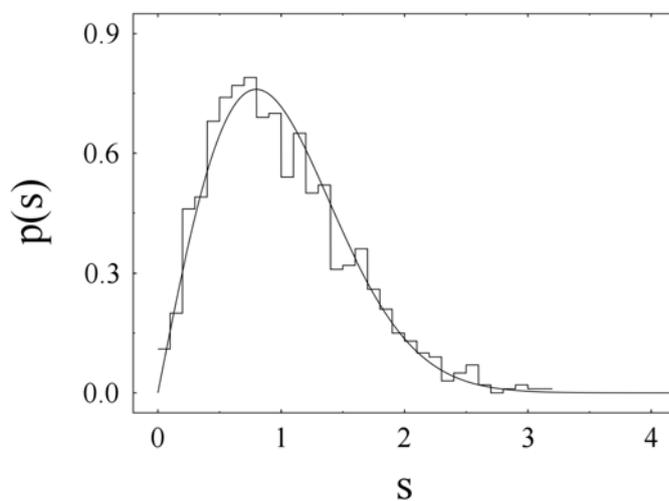


Figura 2.3: Distribución de espaciamiento a primer vecino para los modos electromecánicos de un cristal de cuarzo de forma irregular. Tomada de [9].

Se han clasificado tres ensembles relevantes en mecánica cuántica, definidos en términos de las propiedades de simetría del hamiltoniano [8].

(i) El ensemble ortogonal gaussiano (GOE³). Son sistemas invariantes ante inversión temporal. Para tales sistemas, la matriz hamiltoniana puede ser elegida real y simétrica.

(ii) El ensemble unitario gaussiano (GUE⁴). Son sistemas no invariantes ante reversión temporal. Para tales sistemas, las matrices hamiltonianas son hermitianas y los elementos H_{nm} de las matrices aleatorias son complejos.

(iii) El ensemble simpléctico gaussiano (GSE⁵). Son sistemas invariantes ante inversión temporal, sin simetría rotacional y con espín semi-entero.

En los tres casos la densidad de probabilidad de encontrar una matriz particular está dada por una función $P_D(H)$. Las funciones $P_D(H)$ son invariantes bajo transformaciones ortogonales, unitarias y simplécticas del hamiltoniano, respectivamente.

Para estos ensembles considerados por Wigner, las densidades de probabilidad $P_D(H)$ se escogen con forma gaussiana. Que sean densidades de probabilidad gaussianas se debe a un argumento de máxima entropía en las densidades de probabilidad de los ensembles, además de fijar $\text{tr}(H^2)$ a una constante.

2.2.2. Ensemble de Wishart

Una matriz de Wishart [10] se define como $C = \frac{1}{T}AA^t$ donde A es una matriz de $N \times T$, A^t es la transpuesta de A , y las entradas de A son variables gaussianas independientes reales con promedio cero y varianza $\sigma^2 = 1$. En RMT el ensemble de matrices de Wishart es conocido como ensemble ortogonal de Wishart (WOE) [20]. Si hablamos de series de tiempo, C representa la matriz de correlación calculada sobre las series de tiempo estocásticas de horizonte temporal T para N variables estadísticamente independientes. Esto significa que en promedio C no tiene correlaciones cruzadas.

Para WOE, la Distribución de Probabilidad Conjunta (JPD) de los valores propios de C , λ_j , para $1 \leq j \leq N$, puede ser obtenida usando técnicas desarrolladas para ensembles gaussianos [20]. La JPD de los valores propios es:

$$P(\lambda_1, \dots, \lambda_N) \propto \prod_{j=1}^N w(\lambda_j) \prod_{j>k}^N |\lambda_j - \lambda_k|.$$

donde $w(\lambda_j) = \lambda_j^{[N(\kappa-1)-1]/2} \exp(-N\kappa\lambda_j/2)$ y $\kappa = T/N$.

³ Gaussian orthogonal ensemble

⁴ Gaussian unitary ensemble

⁵ Gaussian symplectic ensemble

A partir de esta función se pueden calcular correlaciones espectrales mediante la función de correlación de n -puntos:

$$R_n(\lambda_1, \dots, \lambda_n) = \frac{N!}{(N-n)!} \int_{-\infty}^{\infty} d\lambda_{n+1} \cdots \int_{-\infty}^{\infty} d\lambda_N P(\lambda_1, \dots, \lambda_N).$$

La función de correlación de n -puntos describe la estadística de $n < N$ valores propios independientemente de los $(N-n)$ valores propios restantes. Por ejemplo, la densidad espectral es la función de un punto.

Para un GOE, la densidad espectral converge al semi-círculo de Wigner para dimensiones grandes. De forma similar, para el WOE, para grades N y T pero finito κ , la densidad espectral está dada por la densidad de Marčenko Pastur [11]:

$$\rho(\lambda) = \kappa \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{2\pi\lambda}, \quad (2.1)$$

donde $\lambda_{\pm} = (\kappa^{-1/2} \pm 1)^2$ son los puntos de los extremos de la densidad. Nótese que para $\kappa < 1$, la densidad $\rho(\lambda)$ está normalizada a κ y no a 1, dado que la matriz C es singular. Por lo tanto, tomando en cuenta los $(N-T)$ ceros, para $\kappa \leq 1$ la densidad toma la forma

$$\rho(\lambda) = \kappa \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{2\pi\lambda} + (1 - \kappa)\delta(\lambda). \quad (2.2)$$

El éxito de la aplicación de RMT en diversos campos se debe a la universalidad de las fluctuaciones espectrales. Pandey y Ghosh [21, 22] han demostrado que los ensembles GOE y WOE comparten las mismas fluctuaciones espectrales, aunque no las mismas densidades.

2.2.3. Ensemble de Wishart correlacionado

Para tomar en cuenta las correlaciones entre las variables se considera $C = \frac{1}{T} \Xi^{1/2} B B^t \Xi^{1/2}$, donde Ξ es una matriz positiva definida y simétrica que define las correlaciones y B tiene las mismas características que A . Dicho ensemble es llamado ensemble ortogonal de Wishart correlacionado (CWOE) [12, 13].

Un tratamiento analítico de este ensemble es más difícil que en el caso no correlacionado. A pesar de ésto, recientemente se ha obtenido la densidad espectral $\rho(\lambda)$ para N y T finitos [23–25], y una relación asintótica para la función de 2-puntos [14]. Las fluctuaciones espectrales para CWOE no son universales siempre, como ocurre en el caso de $\Sigma^2(L)$ [8], para L grande.

Un caso que nos servirá estudiar, y para el cual se tiene solución analítica debido a su sencillez, es el de correlación constante. En este ejemplo, la

matriz Ξ representa la matriz de correlación constante por pares c . Los valores propios de esta matriz están dados por

$$\lambda_1 = Nc + 1 - c \quad \text{y} \quad \lambda_i = 1 - c, \quad (2.3)$$

para $i \neq 1$. Es decir, solo tiene un valor propio que sobresale. En este caso la densidad espectral toma la forma [14]

$$\rho(\lambda) = \rho'(\lambda) + \frac{1}{N} \delta\left(\lambda - \frac{(Nc + 1 - c)(Nc\kappa + 1 - c)}{Nc\kappa}\right), \quad (2.4)$$

$$\rho'(\lambda) = \kappa \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{2\pi(1 - c)\lambda}, \quad (2.5)$$

donde $\lambda_{\pm} = (1 - c)(\kappa^{-1/2} \pm 1)^2$ y $\kappa = T/N$.

Como se puede ver en la función delta de la densidad espectral, el valor propio más grande de la distribución esta en

$$\lambda_1 = \frac{(Nc + 1 - c)(Nc\kappa + 1 - c)}{Nc\kappa} = (Nc + 1 - c) \left(1 + \frac{1 - c}{Tc}\right), \quad (2.6)$$

y éste difiere del valor propio más grande de Ξ (Eq. 2.3) en la dependencia de $(1 - c)/Tc$.

A continuación se hablará de algunas funciones que caracterizan las fluctuaciones en una distribución, por ejemplo, la distribución de espaciamiento a primer vecino (nearest neighbor spacing distribution (NNSD)) y la varianza de numero (level number variance).

2.2.4. Distribución de espaciamiento a primer vecino

Probablemente la distribución de espaciamiento a primer vecino $p(s)$ es la cantidad que más se usa para describir las fluctuaciones a corto alcance en un espectro [9]. Esta función es la densidad de probabilidad que describe el espaciamiento s entre dos niveles contiguos *desdoblados*⁶ ξ_n y ξ_{n+1} . La función $p(s)$ no es una estadística de dos puntos, involucra las correlaciones de n -puntos, $n \leq N$, y su primer momento esta normalizado a la unidad,

$$\int_0^{\infty} p(s) ds = 1 \quad \text{y} \quad \int_0^{\infty} s p(s) ds = 1. \quad (2.7)$$

Para el caso de espectros no-correlacionados o de Poisson se tiene $p(s) = \exp(-s)$, mientras que para el oscilador armónico $p(s) = \delta(s - 1)$. Para los

⁶ La descripción de la técnica de *desdoblamiento* se da en el apéndice A.

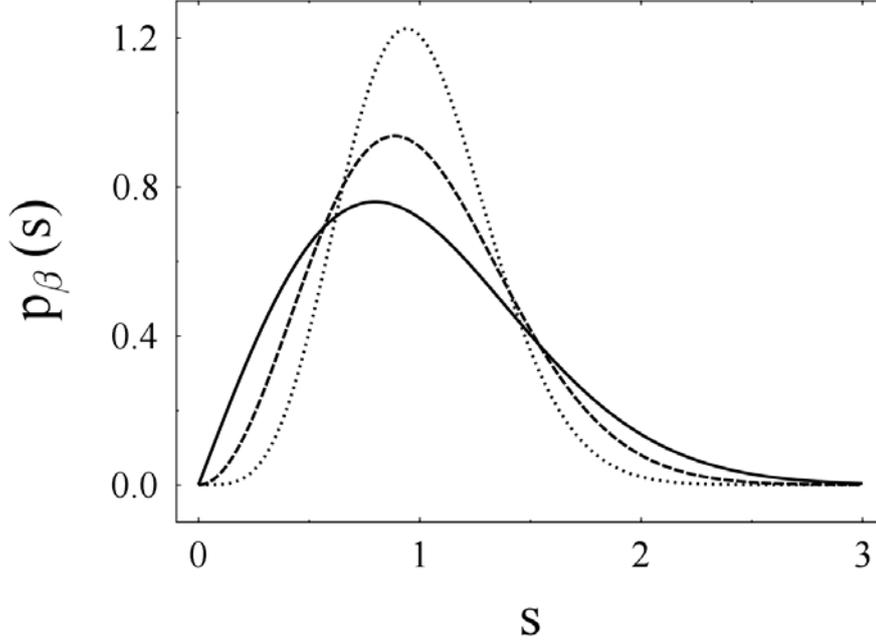


Figura 2.4: La conjetura de Wigner $p_\beta(s)$ para la distribución de espaciamiento a primer vecino. La línea sólida es para GOE, la rayada para GUE, y la punteada para GSE. Tomada de [9].

ensembles gaussianos, el cálculo de $p(s)$ es difícil [20] y el resultado involucra productos infinitos. Sin embargo, una buena aproximación está dada por la conjetura de Wigner, el cual es la solución exacta para un ensemble gaussiano de matrices de 2×2 [26]. La forma más general de la conjetura de Wigner es

$$p_\beta(s) = a_\beta s^\beta \exp(-b_\beta s^2) \quad (2.8)$$

con $\beta = 1$ (GOE), 2 (GUE) y 4 (GSE). La forma de las funciones $p_\beta(s)$ se muestran en la Fig. 2.4. Así, la $p(s)$ para el ensemble GOE está dada por:

$$p_1(s) = p_{GOE}(s) = \frac{\pi}{2} s \exp\left(-\frac{\pi}{4} s^2\right). \quad (2.9)$$

2.2.5. Varianza de número

La distribución de espaciamiento a primer vecino contiene información del espectro a corto alcance. Una forma de medir las correlaciones de largo alcance es mediante la varianza de número, que está definida por

$$\Sigma^2(L) = \langle \hat{\eta}^2(L, \xi_s) \rangle - \langle \hat{\eta}(L, \xi_s) \rangle^2, \quad (2.10)$$

donde $\hat{\eta}(L, \xi_s)$ cuenta el número de niveles en el intervalo $[\xi_s, \xi_s + L]$ de los datos desdoblados y $\langle \cdot \rangle$ denota el promedio sobre todo el intervalo. Por construcción, $\langle \hat{\eta}(L, \xi_s) \rangle = L$. Entonces, en un intervalo de longitud L el promedio de la variación cuadrática es $\Sigma^2(L)$.

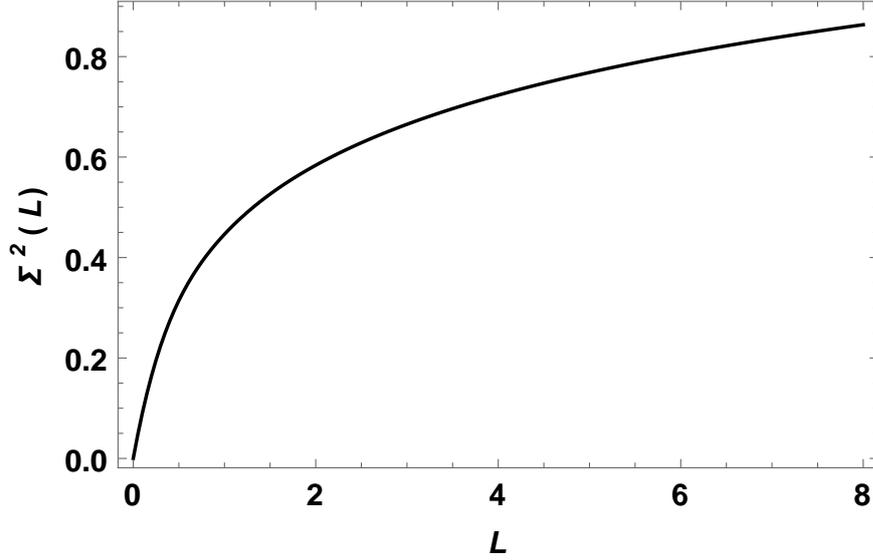


Figura 2.5: Función $\Sigma^2(L)$ para el ensemble GOE. Función 2.13.

La varianza de número solo se relaciona con correlaciones de dos niveles, en la forma siguiente

$$\Sigma^2(L) = L - 2 \int_0^L (L-r) Y_2(r) dr, \quad (2.11)$$

donde $Y_2 = 1 - X_2$; X_2 es la función de correlación de dos puntos para el espectro desdoblado y da la probabilidad de encontrar cualesquiera dos niveles a una distancia r el uno del otro. Para el ensemble GOE está dada por

$$Y_2(r) = s^2(r) + \frac{ds(r)}{dr} \int_r^\infty s(r') dr', \quad (2.12)$$

donde $s(r) = \sin(\pi r)/\pi r$.

Con las ecuaciones anteriores, la solución para $\Sigma^2(L)$ está dada por

$$\Sigma_{GOE}^2(L) = \frac{2}{\pi^2} (\ln(2\pi L) - \cos(2\pi L) + \gamma + 1 - \text{ci}(2\pi L) + \pi^2 L + \text{si}^2(\pi L)/2 - \pi/2 \text{si}(\pi L) - 2\pi L \text{si}(2\pi L)), \quad (2.13)$$

donde $\text{ci}(x) = -\int_x^\infty \cos(t)/t dt$ y $\text{si}(x) = \int_0^x \sin(t)/t dt$. En la Fig. 2.5 se muestra la forma de la función $\Sigma_{GOE}^2(L)$.

Capítulo 3

Descripción del modelo

El análisis de series de tiempo se ha vuelto una herramienta teórica indispensable en la investigación económica y financiera.

Las series de tiempo son observaciones o valores ordenados cronológicamente. Ejemplos de esas variables pueden ser índices de mercados de valores, tasas de desempleo, y precio de acciones. Éstas pueden mostrar distintos patrones, tales como tendencias y periodicidad. Por ejemplo, la producción industrial y el consumo muestran una tendencia ascendente, mientras que el turismo presenta periodicidad.

Debido a esto, aunque el precio de la acción se forma por las fuerzas de oferta y demanda, el rendimiento (o *return*) a menudo es la variable que se analiza. Su ventaja es que algunas propiedades estadísticas, tal como estacionariedad, pueden ser más aplicables a los rendimientos que a los precios [27], puesto que calcular rendimientos es una forma de quitar tendencias.

Para una acción i que tiene un precio $P_i(t)$ al tiempo t , el rendimiento se define como

$$r_i(t) = \frac{P_i(t + \Delta t)}{P_i(t)} - 1, \quad (3.1)$$

donde Δt representa una unidad de tiempo en la serie de tiempo. Esta unidad de tiempo pueden ser segundos, minutos, días, etc... En este trabajo la unidad de tiempo será de un día laboral¹.

Una herramienta común en el estudio de la interdependencia entre rendimientos de diversos bienes, entre otros acciones, es la matriz de correlación C de rendimientos:

$$C = \text{diag}(\Sigma)^{-1/2} \Sigma \text{diag}(\Sigma)^{-1/2}, \quad (3.2)$$

¹ En el siguiente capítulo se explican con más detalle las características de los datos

donde $\text{diag}(\Sigma)$ es una matriz diagonal con la diagonal de la matriz de covarianza Σ , que se define como:

$$\Sigma_{ij} = \text{cov}(r_i(t), r_j(t)) = E[(r_i(t) - \mu_i)(r_j(t) - \mu_j)], \quad (3.3)$$

donde μ_i es el promedio de la serie de tiempo i .

Si las series de tiempo son previamente normalizadas, es decir, si $A_i(t) = (r_i(t) - \mu_i)/\sigma_i$, entonces la matriz de correlaciones se puede escribir como:

$$C = \frac{1}{T} AA^t. \quad (3.4)$$

En trabajos anteriores se ha analizado la matriz de correlación C para datos del mercado de valores de EE.UU.. De los cuales, en algunos se ha hecho énfasis en la aplicación de RMT [3–7], principalmente en el análisis del espectro de valores propios de las matrices de correlación.

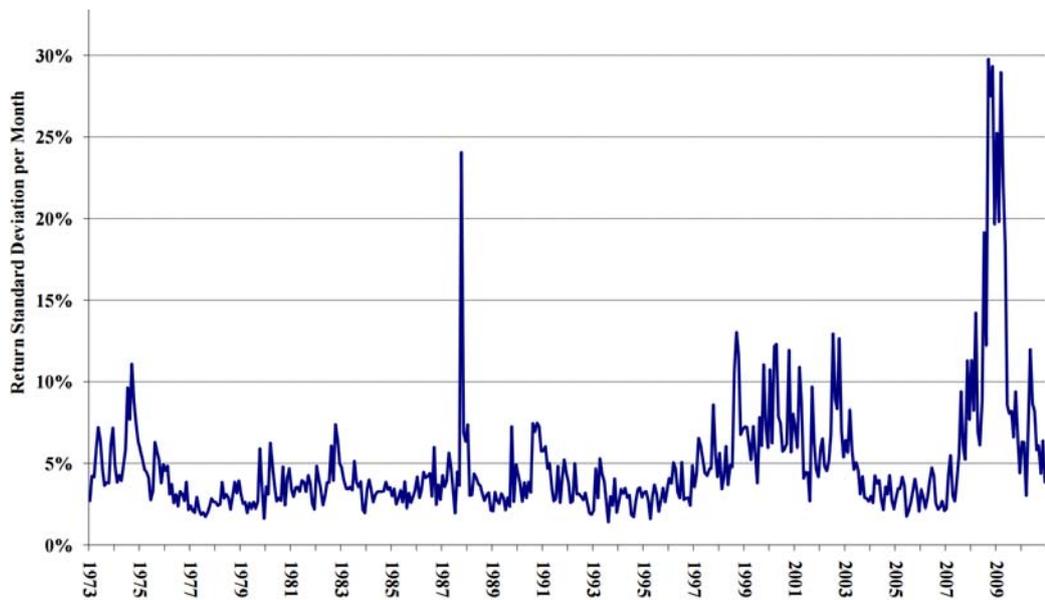


Figura 3.1: Volatilidad de los rendimientos en Datastream US Financial Index. Tomada de [28].

Los valores propios λ_i y vectores propios e_i de C tienen una interpretación económica. Al valor propio más grande se le identifica con el movimiento del mercado, y a los siguientes valores propios que se desvían del caso aleatorio se les identifica con sectores industriales [3]; los valores propios restantes básicamente siguen una distribución de Marčenko Pastur [5]. Esto es importante

porque la matriz de covarianza se usa para hacer optimización de portafolio, y que la matriz de correlación tenga las características mencionadas en el párrafo anterior genera dudas en dicha optimización. Por otro lado, Münnix et al. [2] proponen un enfoque en el que detectan estados del mercado, partiendo de la aglomeración de matrices de correlación en el tiempo.

A la medida de la dispersión de los rendimientos de un instrumento financiero en un período dado se le llama volatilidad [29], y una forma de medirla es mediante la desviación estándar de los rendimientos. A lo largo de los años se ha encontrado que una volatilidad grande está relacionada con una crisis financiera [28] (por ejemplo Fig. 3.1), y en años más recientes se han relacionado correlaciones promedio grandes con volatilidad alta [30–36]. Además, se relaciona una correlación promedio grande con crisis financieras y, como se verá adelante, estas correlaciones grandes derivan en un valor propio que sobresale notablemente del resto, el cual representa el movimiento colectivo del mercado.

Hay suficiente evidencia de que la volatilidad cambia en el tiempo [37], y no solo cambia sino que presenta *volatility clustering* [1], esto quiere decir que grandes cambios tienden a preceder grandes cambios (de cualquier signo) y lo mismo ocurre con pequeños cambios [38]. Por estas razones, la serie de tiempo de rendimientos no es estacionaria en el sentido estricto². Aunque una serie de tiempo suficientemente larga puede ser estacionaria asintóticamente en el mejor de los casos [40].

Motivados por el hecho de que estas series de tiempo presentan no estacionariedad, a continuación se presenta un modelo para estudiar series de tiempo cortas, con la asunción de que para periodos de tiempo cortos podemos encontrar más estacionariedad, y extraer propiedades estadísticas del espectro de valores propios de las matrices de correlación.

3.1. Modelo

El objetivo de este trabajo es analizar series de tiempo financieras cortas obtenidas a partir de series largas no-estacionarias. Con este fin, se construye un ensemble de matrices de correlación a partir de un conjunto de series de tiempo cortas. Este método se explica a continuación.

² Un proceso estocástico $x(t)$ es estacionario si su densidad de probabilidad (PDF) $P[x(t)]$ es invariante bajo traslaciones temporales. Esta es una definición rigurosa, y es llamada *estacionariedad en el sentido estricto*. Existen definiciones menos restrictivas de un proceso estocástico estacionario [39]. Entre estas se encuentra la de un *proceso estocástico asintóticamente estacionario*, estos son observados cuando la estadística de las variables aleatorias $x(t_1 + c), \dots, x(t_n + c)$ no depende de c si c es grande.

Sea A una matriz de series de tiempo normalizadas de dimensión $N \times T$,³ es decir, el renglón A_i representa la serie de tiempo i donde $A_{ij} \rightarrow (A_{ij} - \langle A_i \rangle) / \sigma_i$. En este caso tomamos $T \ll N$, por lo que la matriz de correlación es singular y tiene $N - T + 1$ ceros como valores propios, he aquí el por qué no se puede analizar el espectro de valores propios, si se calcularan los valores propios se tendría pocos valores diferentes de cero.

Ésta es la razón fundamental por la que se crea el ensemble de matrices de correlación pequeñas, ya que, al calcular los valores propios de cada matriz se tendrán muchos más valores propios, con los que se podrá hacer estadística.

En el análisis que desarrollaremos la matriz A es una matriz de rendimientos, es decir, el renglón A_i representa la serie de tiempo de los rendimientos normalizados para la acción i . Decimos que son series de tiempo cortas porque tomamos una época de aproximadamente dos meses caracterizada por el tiempo inicial τ del horizonte que se considera, por lo que las matrices varían con el tiempo. Esto se tratará con más detalle en la sección de resultados donde se hablará de los datos que se analizan.

Si se toma una matriz A para un tiempo dado, el ensemble de matrices de correlación $\{\mathcal{C}_\alpha\}$ se construye de la siguiente forma:

1. Para crear un miembro \mathcal{C}_α del ensemble se escogen $n < T$ series de tiempo diferentes de A , y con estas series de tiempo se crea una matriz de correlación. A continuación se muestra un esquema

$$A \rightarrow \mathcal{A}_\alpha \rightarrow \mathcal{C}_\alpha = \frac{1}{T} \mathcal{A}_\alpha \mathcal{A}_\alpha^t,$$

donde \mathcal{A}_α representa la matriz de rendimientos creada con las n series de tiempo elegidas y el subíndice α se usa para distinguir el conjunto particular seleccionado.

2. Se crea un conjunto de diferentes matrices de correlación \mathcal{C}_α , este conjunto representa nuestro ensemble. Para que sean diferentes las matrices de correlación se tiene que escoger un conjunto diferente de series de rendimientos para cada miembro del ensemble, es decir, α es diferente para cada miembro,

$$\{\mathcal{C}_\alpha\} = \{\mathcal{C}_{\alpha_1}, \mathcal{C}_{\alpha_2}, \dots\}.$$

En el paso número 1, se tienen que escoger $n < T$ series de tiempo de A , pero ¿cómo se escogen?

Estudiaremos dos maneras de elegir estas series de tiempo:

³ Las unidades de tiempo de T son días laborales.

- Aleatoria: Se eligen n series de tiempo de forma aleatoria, tal que no se repitan las series de tiempo dentro de una matriz.
- Por sector industrial⁴: Se eligen n series de tiempo tomando en consideración cuántas acciones corresponden a los diferentes sectores. Si se tienen N_s series de tiempo del sector s , se tomarán $n_s = n N_s/N$ series de tiempo de dicho sector, de tal forma que $\sum_s n_s = n$. Y dentro de cada sector se escogen las n_s acciones de forma aleatoria.

La elección de series de tiempo equivale a elegir renglones y columnas de la matriz de correlación C . Esto es debido a que las entradas de una matriz C_α son elementos de la matriz C . Así, se puede pensar a la matriz C como una matriz de referencia para crear el ensemble.

3.2. Correlaciones parciales

Recientemente se ha utilizado como método de estudio de series de tiempo la correlación parcial [15]. La correlación parcial es una herramienta que se usa para saber si la correlación entre dos series de tiempo, por ejemplo rendimientos, es el resultado de su correlación con una tercera. Por ejemplo, supongamos que tenemos tres acciones A , B y C , y que encontramos una correlación significativa entre cada par. Si sospecháramos que la correlación entre A y B es el resultado de su correlación individual con C , es decir A - C y B - C , podríamos remover la relación entre A y C , y entre B y C . Ahora calculamos nuevamente la correlación entre A y B después de remover el efecto de C , la cual es la correlación parcial. Si la correlación parcial es significativamente más pequeña que la original, entonces, la correlación se debía principalmente a su correlación con C .

El uso de las correlaciones parciales para estudiar sistemas complejos es cada vez más popular. Se ha implementado en el estudio de redes de genes [41–43], y además se ha usado en el mercado financiero [44–46].

De esta forma la correlación entre A y B dado C , esta dada por

$$c(A, B : C) = \frac{c(A, B) - c(A, C) c(B, C)}{\sqrt{[1 - c^2(A, C)][1 - c^2(B, C)]}}, \quad (3.5)$$

⁴ La clasificación GICS (Global Industry Classification Standard) fue propuesta por Standard & Poor's y MSCI Barra en 1999, y ahora es reconocida internacionalmente. A las compañías se les asigna un sector dependiendo de cuál sea su principal actividad económica. Hay 10 sectores: Energy, Materials, Industrials, Consumer Discretionary, Consumer Staples, Health Care, Financials, Information Technology, Telecommunication Services y Utilities.

donde $c(i, j)$ representa la correlación de Pearson entre i y j . En nuestro caso, se calculará la correlación parcial entre las acciones i y j dado r_{sp} , donde r_{sp} es la serie de tiempo de rendimientos promedio durante ese intervalo de tiempo.

En el análisis se toman tanto la matriz de correlación ordinaria como la matriz de correlación parcial como matriz de referencia para crear el ensemble. Se incluye la correlación parcial porque puede ser que la relación intrínseca entre i y j no esté bien reflejada por la correlación ordinaria, debido a que el efecto del movimiento promedio del mercado es muy grande en las matrices de correlación [5].

Capítulo 4

Resultados

El conjunto de acciones que se analizaron corresponden a $N = 293$ acciones del índice S&P 500 durante el período de tiempo entre 1992 y principios de 2015. Las empresas escogidas han tenido una participación ininterrumpida en el índice S&P 500 durante el periodo estudiado y corresponden a las estudiadas en [2], salvo algunas que se sacaron por no estar presentes hasta 2015. Las series de tiempo de precios de las acciones corresponden a precios de cierre diarios, por lo que $\Delta t = 1$ día¹. Estudiamos estos precios porque estamos interesados en tendencias de larga escala y es fácil acceder² a este tipo de datos a través de Yahoo! Finance.

A partir de los precios se calculan las series de tiempo de rendimientos para cada acción. Los rendimientos a su vez se dividen en series más pequeñas, de aproximadamente dos meses de días laborales, $T = 44$. Estas series de tiempo cortas son las que se analizan con el método propuesto. Además, las series de tiempo que corresponden a dos tiempos contiguos se traslapan en³ $3/4$, esto se hace para obtener la mayor información conforme movemos la ventana de tiempo. En la Fig. 4.1 se muestra un esquema de este proceso.

Para crear el ensemble se siguen los pasos que se explicaron en el capítulo anterior y se toman $n = 35$ series de tiempo, por lo que la dimensión de la matriz de correlación \mathcal{C}_α es 35×35 . Entonces, para cada miembro del ensemble se tendrán 35 valores propios, todos diferentes de cero, por consiguiente, habrá un ensemble para el valor propio más grande, para el segundo, y lo

¹ Solo se toman los días que abre la bolsa, por lo tanto no se toman fines de semana, días festivos y acontecimientos imprevistos (ej. el atentado a las torres gemelas de NY). En [2] quitan los datos del lunes porque es un salto de 36 horas (viernes-lunes), mientras que los días restantes de la semana son saltos de 24 horas. Nosotros tomaremos todos los datos y los trataremos de igual manera.

² En cambio, los precios a escala de operaciones (intraday) son caros y/o de difícil acceso.

³ Lo hacemos de este modo aunque es más común hacerlos con $1/2$.

mismo para el resto. El tamaño del ensemble es de 5000 elementos.

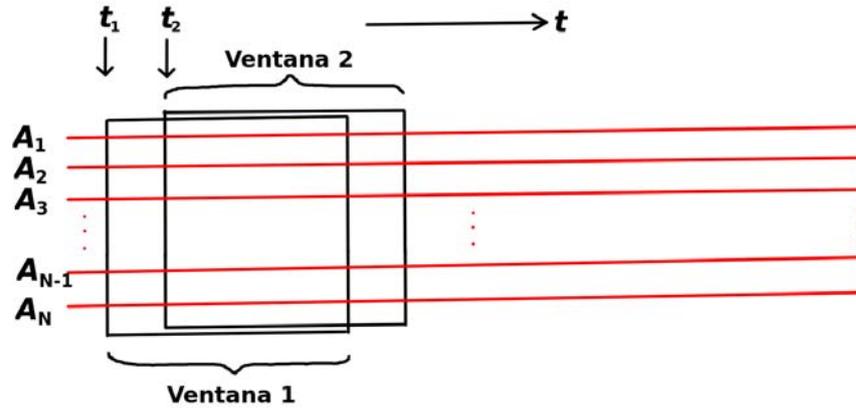


Figura 4.1: Esquema de como se cortan las series de tiempo en series de tiempo cortas de T días. Además las ventanas contiguas se traslapan $3/4$.

Para crear la matriz de correlación parcial se toman correlaciones parciales entre las acciones i y j dado el rendimiento promedio r_{sp} de las 293 acciones, es decir, $c(i, j : r_{sp})$.

4.1. Distribución de valores propios

Como ya se explico antes, con el método se obtiene un ensemble de valores para cada uno de los 35 valores propios. En las siguientes apartados se estudiara como son las densidades espectrales.

Usualmente la densidad espectral se divide principalmente en dos partes, la primera es el denominado *bulk*, que está formado por la mayoría de los valores propios excepto los más grandes, y la segunda está dada por unos cuantos de los valores propios más grandes, que corresponden a los sectores industriales. Debido al método que estamos usando, solo el valor propio más grande se separa notablemente del resto para correlaciones ordinarias, y el resto permanece cerca, por lo que llamaremos bulk a la distribución formada por todos los valores propios salvo el mayor. Por lo tanto, el bulk cuenta con $34 \times 5000 = 170000$ valores propios, y la distribución para el valor propio más grande con 5000.

Cabe destacar que esto ocurre cuando se usan correlaciones ordinarias, pero cuando se trabaja con correlaciones parciales el valor propio más grande no se separa del resto.

A pesar de esto, sí se observa que algunos valores propios, a demás del mayor, se separan un poco. Esto se muestra en la sección 4.1.3.

4.1.1. Bulk

Debido a que la distribución del valor propio más grande es la única que se separa de forma importante, el bulk se puede intentar comparar con la distribución del modelo de Wishart de correlación constante.

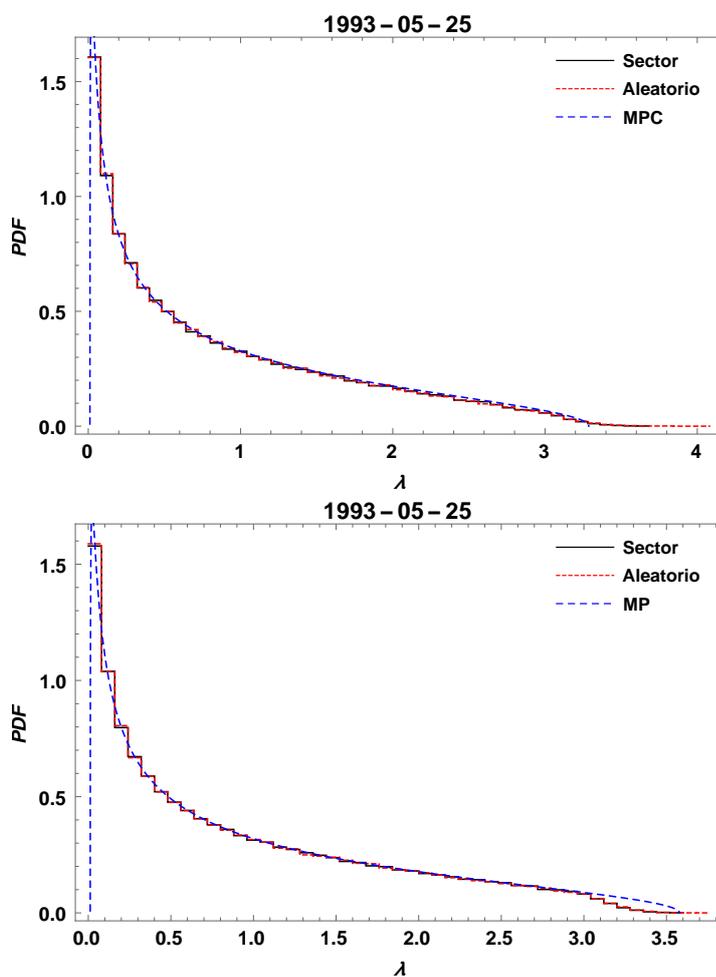


Figura 4.2: Bulk para elección aleatoria, sectores, MP y MPC calculado con c_{est} , para correlaciones ordinarias (arriba) y cuando se usa correlaciones parciales (abajo). La distribución para el selección por sectores y aleatoria son indistinguibles.

En la Ec. 2.6 se relaciona el valor propio mayor con la correlación cons-

tante. De esta ecuación podemos poner a c en términos de λ_1 :

$$c = \frac{b + \sqrt{b^2 - 4a}}{2a}, \quad (4.1)$$

donde $b = T(\lambda_1 - 1) - N + 2$ y $a = (T - 1)(N - 1)$.

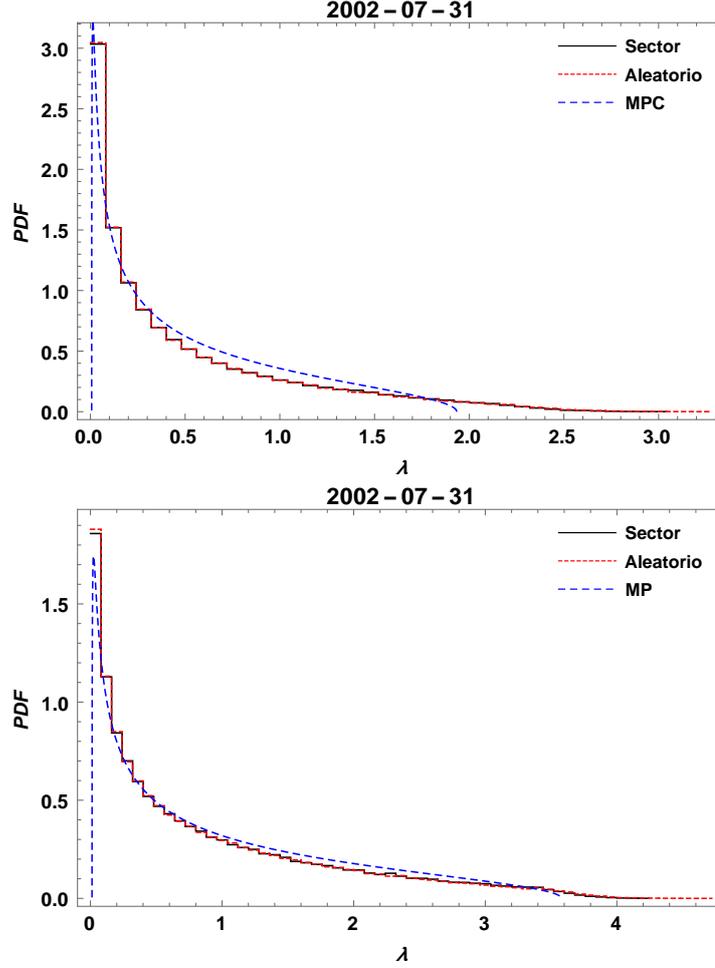


Figura 4.3: Bulk para elección aleatoria, sectores, MP y MPC calculado con c_{est} , para correlaciones ordinarias (arriba) y cuando se usa correlaciones parciales (abajo).

Para hacer una estimación de la correlación c para nuestro ensemble, dado este modelo de correlación constante, hacemos $N \rightarrow n$ y $\lambda_1 \rightarrow \langle \lambda_1 \rangle$, donde éste último represente el valor promedio de la distribución para el valor propio más grande, así

$$c_{est} = \frac{b' + \sqrt{b'^2 - 4a'}}{2a'}, \quad (4.2)$$

donde $b' = T(\langle \lambda_1 \rangle - 1) - n + 2$ y $a' = (T - 1)(n - 1)$.

Para el bulk con correlaciones ordinarias se ajusta la distribución de Marčenko Pastur correlacionado (MPC) dada por Ec. 2.5, tomando la correlación estimada de Ec. 4.2, y para las correlaciones parciales se ajusta la distribución de Marčenko Pastur (MP) dada por Ec. 2.1. Algunos ejemplos de estas distribuciones, para diferentes tiempos, se muestran en las figuras 4.2, 4.3 y 4.4. Las curvas de arriba (en la figura) muestran el bulk para correlaciones ordinarias y las de abajo para correlaciones parciales, pero para el mismo tiempo.

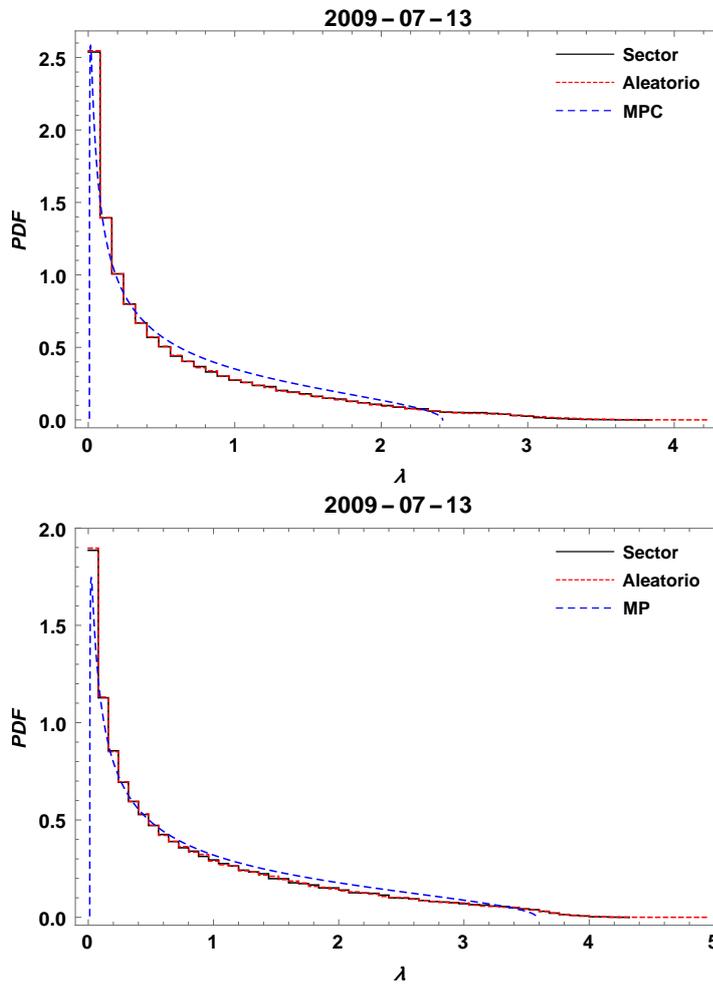


Figura 4.4: Bulk para elección aleatoria, sectores, MP y MPC calculado con c_{est} , para correlaciones ordinarias (arriba) y cuando se usa correlaciones parciales (abajo).

En el caso de correlaciones ordinarias, la distribución de correlación cons-

tante se acerca más al bulk que la distribución de Marčenko Pastur, pero el bulk para correlaciones parciales se ajusta mejor MP, lo cual concuerda con el hecho de haber removido el efecto promedio del movimiento del mercado, que se identifica con el valor propio más grande.

4.1.2. Valor propio más grande

En general, cuando se analizan los valores propios derivados de las correlaciones de Pearson, el valor propio más grande se separa claramente para la mayoría de los tiempos, como se muestra en la figura izquierda de 4.5, donde $\langle \lambda_1 \rangle_R$ es el valor promedio de la distribución del valor propio más grande para el ensemble aleatorio. El valor promedio de la distribución del máximo valor propio se define como

$$\langle \lambda_1 \rangle_Q = \frac{1}{l} \sum_{\alpha \in Q} \lambda_{1,\alpha}, \quad (4.3)$$

donde $\lambda_{1,\alpha}$ es el valor propio más grande de la matriz de correlación \mathcal{C}_α , Q puede ser R o S , R representa el ensemble aleatorio y S por sectores, y l en tamaño del ensemble.

Cuando se estudian las correlaciones parciales no se observa esto, los valores propios más grandes permanecen pequeños. Los dos picos sobresalientes que se observan para correlaciones parciales corresponden con dos crisis financieras importantes: la llamada dot-com bubble alrededor de 2001-2002 y la crisis hipotecaria del 2008. Por lo tanto, el valor propio más grande para las correlaciones parciales nos da cierta información de la situación financiera del mercado.

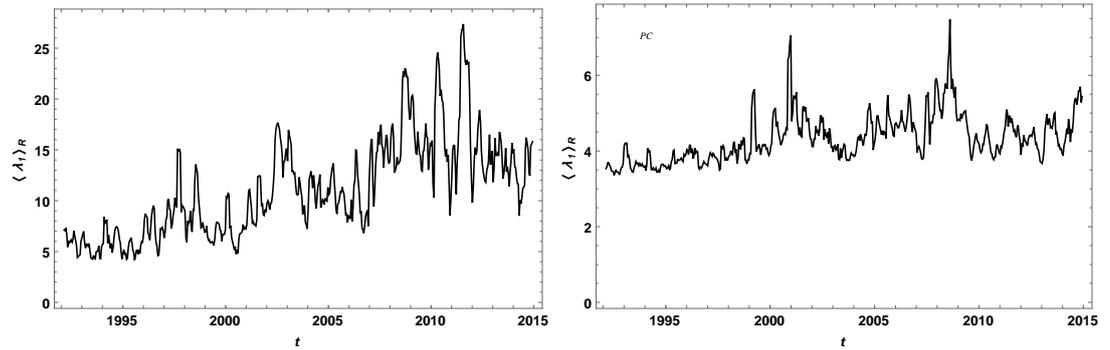


Figura 4.5: Comparación entre el valor promedio del valor propio más grande $\langle \lambda_1 \rangle_R$ cuando se utiliza correlación ordinaria (izquierda) y cuando solo se trabaja con correlación parcial (derecha).

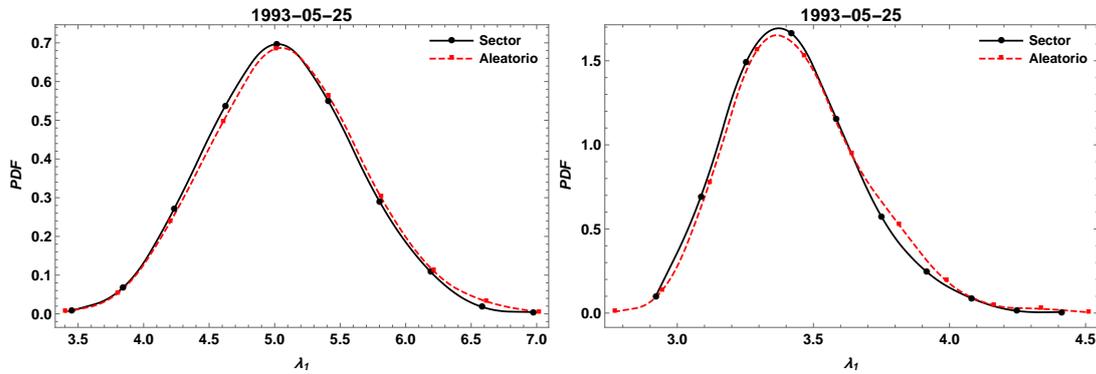


Figura 4.6: Distribución de λ_1 para $t = 1993 - 05 - 25$, tanto con elección aleatoria y por sectores. Correlaciones ordinarias (izquierda) y correlaciones parciales (derecha).

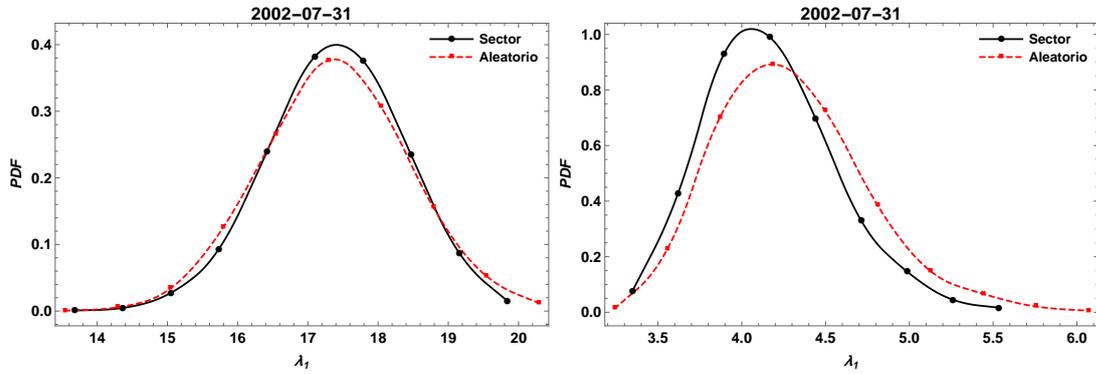


Figura 4.7: Distribución de λ_1 para $t = 2002 - 07 - 31$, tanto con elección aleatoria y por sectores. Correlaciones ordinarias (izquierda) y correlaciones parciales (derecha).

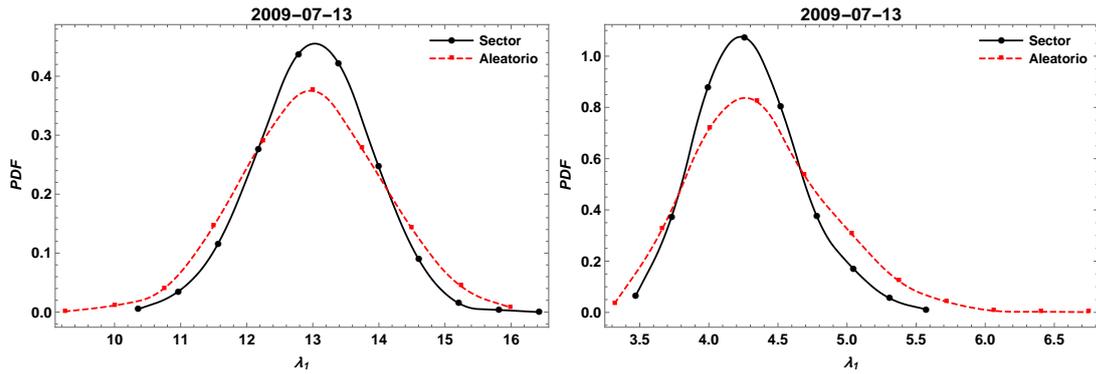


Figura 4.8: Distribución de λ_1 para $t = 2009 - 07 - 13$, tanto con elección aleatoria y por sectores. Correlaciones ordinarias (izquierda) y correlaciones parciales (derecha).

La forma de las distribuciones para el valor propio más grande varía con el tiempo como se muestra en las figuras 4.6, 4.7 y 4.8. Al igual que antes, las figuras del lado izquierdo corresponde al análisis con correlación ordinaria y las de la derecha correlaciones parciales. Cuando se usa correlación parcial el valor propio más grande no se aleja del bulk y mantiene valores pequeños a diferencia del caso con correlaciones ordinarias. Las líneas continuas que unen a los puntos representan interpolaciones de segundo orden entre los puntos, aunque para estas distribuciones se ajusta perfectamente una expansión de Edgeworth⁴ de segundo orden, algunos ejemplos se muestran en las figuras 4.9 y 4.10.

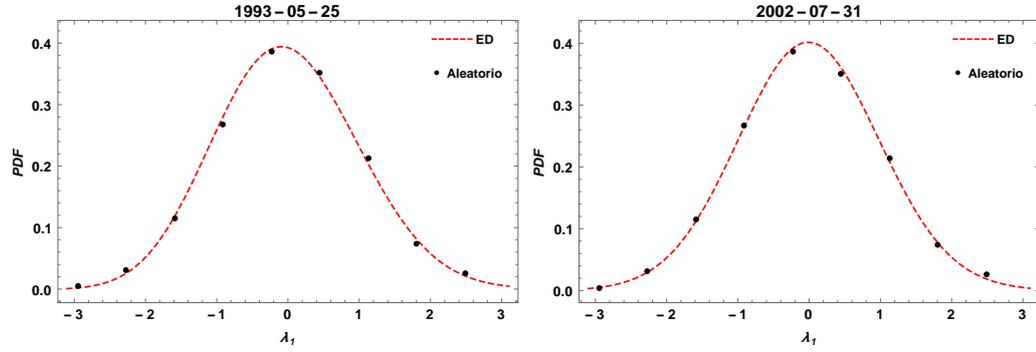


Figura 4.9: Las curvas son las series de Edgeworth a segundo orden para las distribuciones de λ_1 del ensemble aleatorio y diferentes tiempos. Los puntos son los datos reales de las distribuciones después de quitarles el promedio y la varianza.

Algo que se tiene que tomar en cuenta es que la suma de los valores propios para cada matriz del ensemble es $\sum_{j=1}^n \lambda_j = n$. En nuestro caso $n = 35$, y como se observa en la Fig. 4.5, el valor $\langle \lambda_1 \rangle$ llega a 27, lo cual es muy grande. Además, debido a esta restricción en la suma de los valores propios, cuando hay un sólo valor propio que se aleja mucho, el resto se hacen más pequeños. Lo cual se observa en los datos analizados. Por ejemplo, en la gráfica izquierda de Fig. 4.7 el valor promedio está al rededor de 17, lo cual hace que el bulk izquierdo de la Fig. 4.3 se acerque al origen.

⁴ La idea de esta expansión es expresar una aproximación de la función característica de la distribución en términos de su kurtosis κ_4 , skewness κ_3 y una distribución normal. La función característica es

$$g(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \left(1 + \frac{\kappa_3 H_3(x)}{6} + \frac{\kappa_4 H_4(x)}{24} + \frac{\kappa_3^2 H_6(x)}{72}\right),$$

donde $x = (\lambda_1 - \langle \lambda_1 \rangle) / \sigma_{\lambda_1}$.

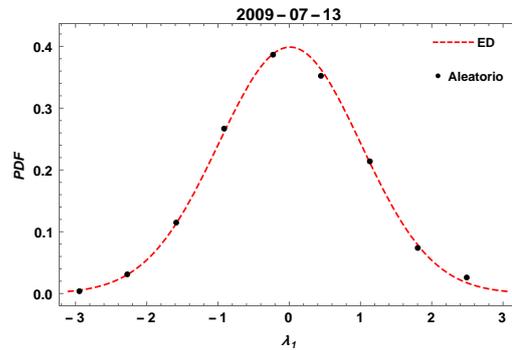


Figura 4.10: La curva es la serie de Edgeworth a segundo orden para la distribución de λ_1 del ensemble aleatorio. Los puntos son los datos reales de la distribución después de quitarles el promedio y la varianza.

4.1.3. Efecto de la estructura de la matriz de correlación

Para saber como se separan los valores propios entre si, para un tiempo dado, estudiamos el valor promedio de sus distribuciones⁵. Una forma de verlo es haciendo una gráfica zipf, como la que se muestra en la Fig. 4.11. En esta figura se observa que, además del primero, el segundo valor propio también se separa del resto para correlaciones ordinarias y se compara con lo obtenido para el caso de datos sintéticos⁶ no correlacionados.

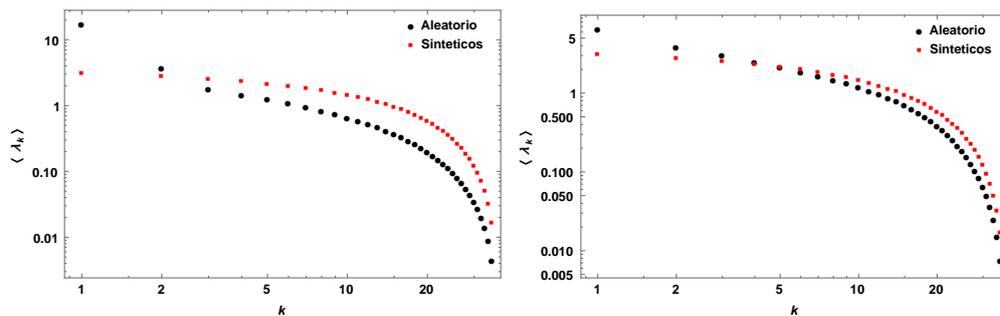


Figura 4.11: Gráfica zipf para el valor promedio de los valores propios con elección aleatoria, $t = 2008-07-25$, correlaciones ordinarias (izquierda) y correlaciones parciales (derecha). Se compara con lo obtenido para datos sintéticos sin correlaciones.

⁵ Se muestran resultados para selección Aleatoria pero lo mismo es aplicable para Sectores.

⁶ Se crean el mismo número de series de tiempo N sin correlaciones entre ellas y se aplica el método para obtener un ensemble de matrices de correlación.

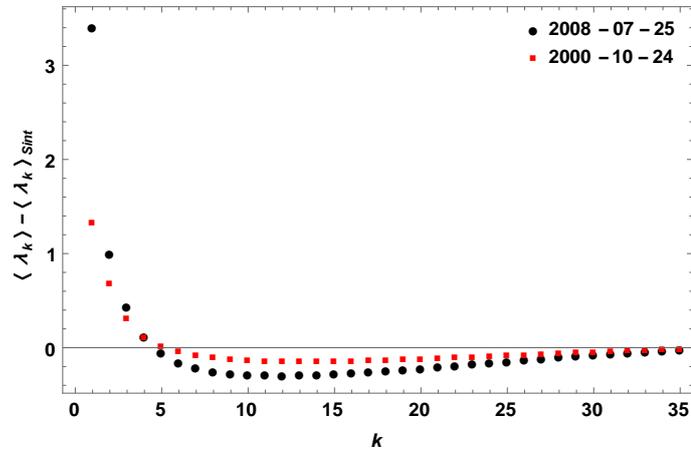


Figura 4.12: Diferencia entre los valores propios obtenidos para el mercado con correlación parcial y para los datos sintéticos. Se muestran dos tiempos distintos y elección Aleatoria.

Si tomamos la diferencia entre los promedios de las distribuciones para el mercado con correlaciones parciales y los datos sintéticos, se obtienen curvas como las que se muestran en la Fig. 4.12, donde se observa que mientras más grande sea la diferencia para los primeros valores propios, la curva alcanzará un valor más negativo en los valores propios medianos.

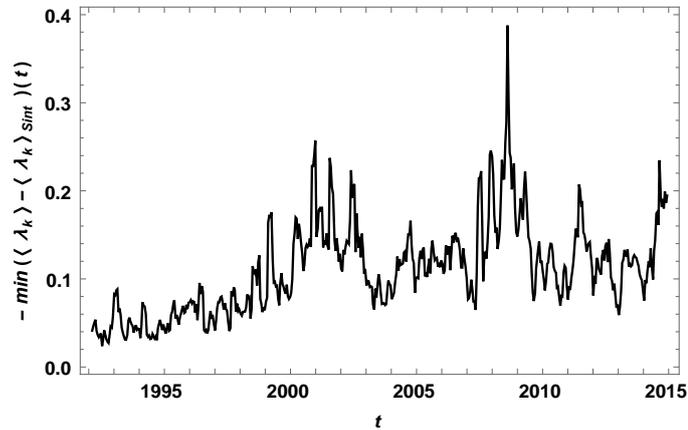


Figura 4.13: Evolución del mínimo de la diferencia entre los valores propios promedio del mercado con correlación parcial y aquellos para datos sintéticos sin correlaciones. Para selección Aleatoria.

Como una forma de cuantificar el efecto de que tanto se alejan los primeros

valores propios del caso no correlacionado, calculamos el valor mínimo de la curvas de la Fig. 4.12. La evolución de este mínimo se muestra en la Fig. 4.13 y se puede ver que en las dos crisis financieras más sobresalientes este valor es pequeño.

La curva de la Fig. 4.13 se obtuvo con la diferencia entre los datos del mercado para correlaciones parciales y los datos sintéticos sin correlaciones. Al quitar la tendencia del mercado estamos quitando las correlaciones promedio, entonces, lo que queda con las correlaciones que reflejan la estructura de la matriz de correlaciones, es por esto que mínimos grandes en 4.13 están relacionados con más valores propios que se separan del bulk para correlaciones ordinarias.

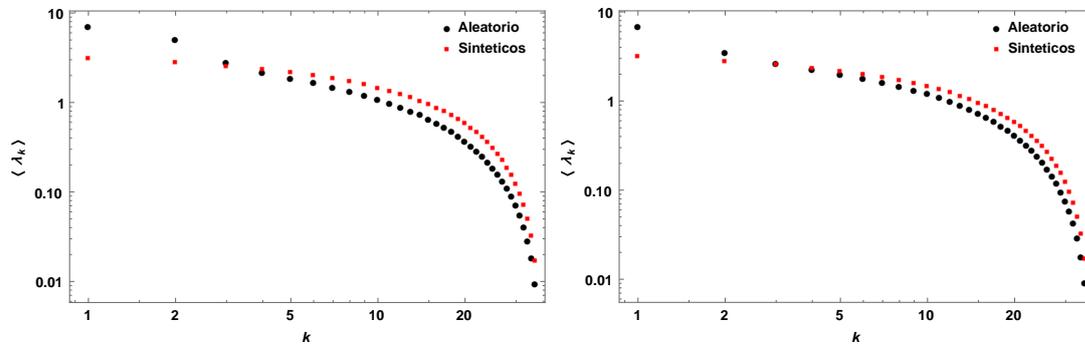


Figura 4.14: Gráfica zipf para el valor promedio de los valores propios con elección Aleatoria, $t = 2000-12-27$, correlaciones ordinarias (izquierda) y correlaciones parciales (derecha). Se compara con lo obtenido para datos sintéticos sin correlaciones.

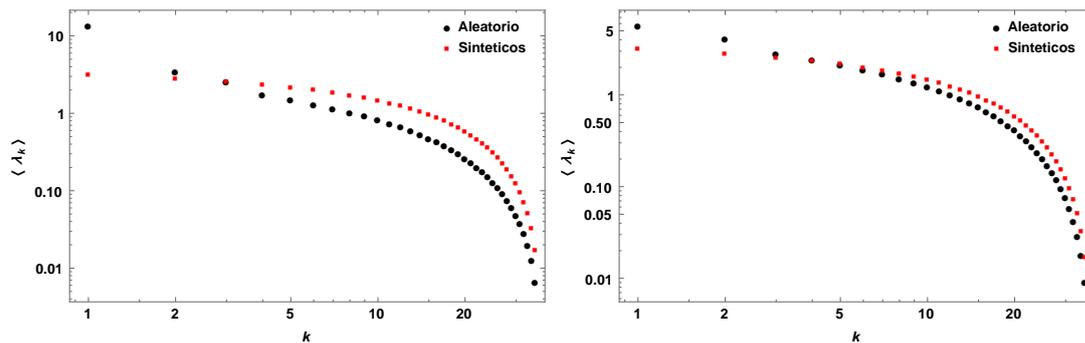


Figura 4.15: Gráfica zipf para el valor promedio de los valores propios con elección aleatoria, $t = 2007-12-31$, correlaciones ordinarias (izquierda) y correlaciones parciales (derecha). Se compara con lo obtenido para datos sintéticos sin correlaciones.

En las figuras 4.11, 4.14 y 4.15 se muestran ejemplos de lo anterior. Donde

se observa que para correlaciones ordinarias el segundo valor propio también se separa (o desde el tercero como en 4.15).

Mínimos pequeños en 4.13 solo muestran que se separa un valor propio para correlaciones ordinarias, como en el caso de las figuras 4.16 y 4.17. Para estos casos el espectro para correlaciones parciales se parece al caso sin correlaciones.

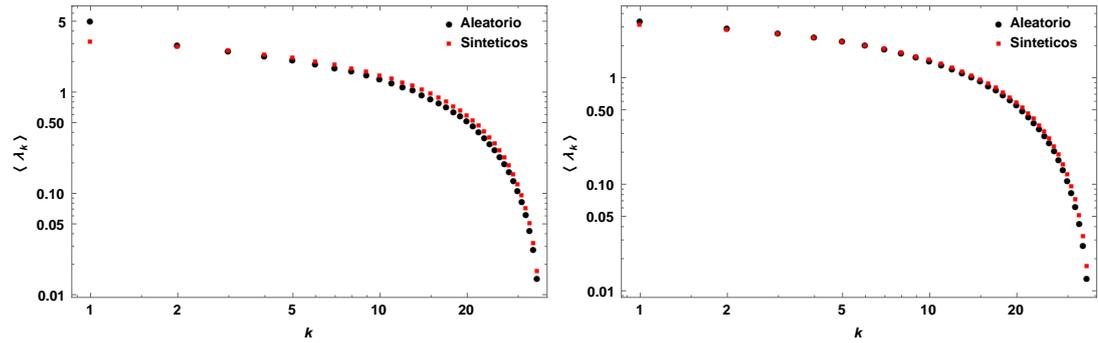


Figura 4.16: Gráfica zipf para el valor promedio de los valores propios con elección aleatoria, $t = 1993-05-25$, correlaciones ordinarias (izquierda) y correlaciones parciales (derecha). Se compara con lo obtenido para datos sintéticos sin correlaciones.

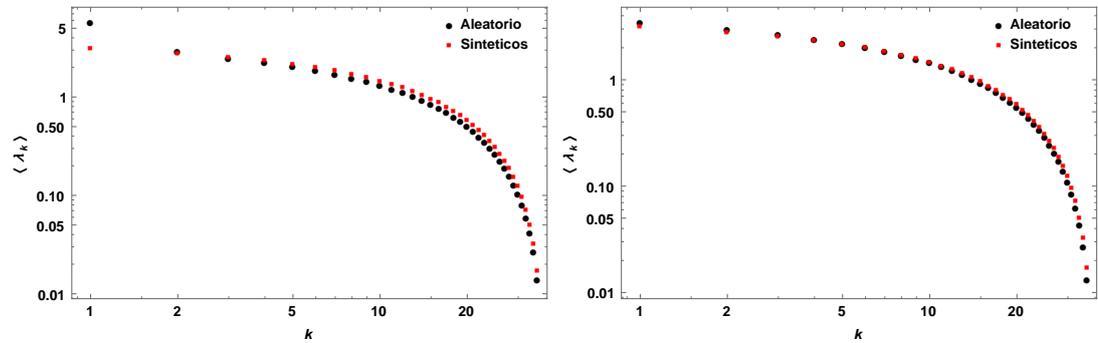


Figura 4.17: Gráfica zipf para el valor promedio de los valores propios con elección aleatoria, $t = 1994-05-10$, correlaciones ordinarias (izquierda) y correlaciones parciales (derecha). Se compara con lo obtenido para datos sintéticos sin correlaciones.

En las figuras 4.18 y 4.19 se muestra el perfil de correlaciones de la matriz de correlación promedio para un tiempo dado, tanto para correlación ordinaria como para correlación parcial. En ambas figuras se observa que los dos tipos de correlación comparten una estructura. La Fig. 4.18 representa un caso en el que se tiene más de un valor propio que se separa del bulk y la Fig. 4.19 es un ejemplo en donde únicamente un valor propio se separa.

Esto es sobresaliente porque generalmente se identifica al valor propio más grande con el estado del mercado pero esto sugiere que la estructura de las correlaciones es igualmente importante.

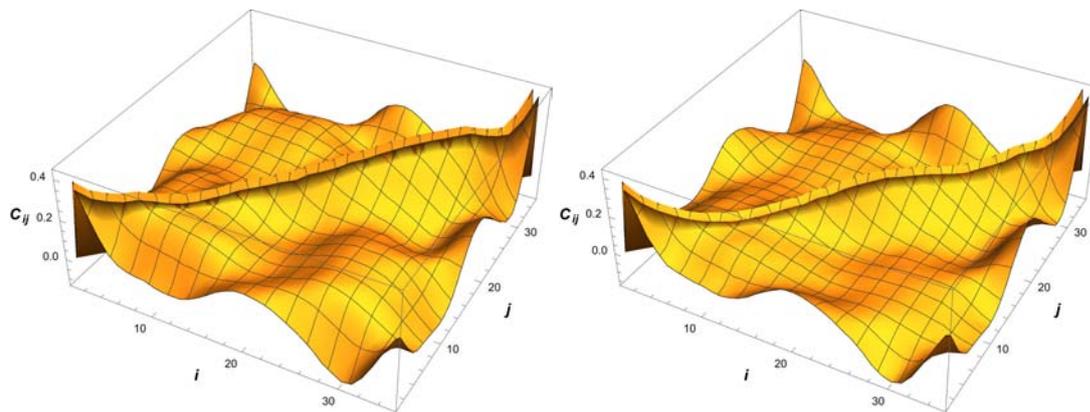


Figura 4.18: Estructura de la matriz de correlación promedio sin diagonal para $t = 2000-12-27$, para correlación ordinaria (izquierda) y correlación parcial (derecha). Para elección aleatoria.

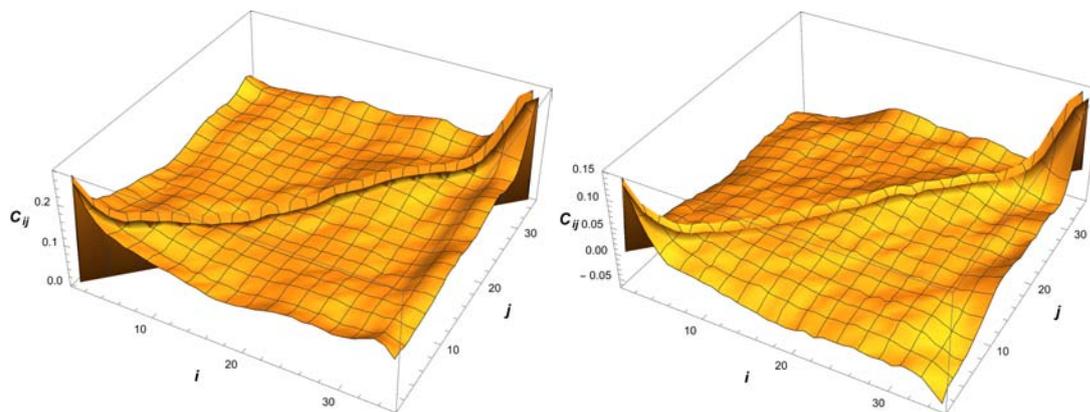


Figura 4.19: Estructura de la matriz de correlación promedio sin diagonal para $t = 1994-01-18$, para correlación ordinaria (izquierda) y correlación parcial (derecha). Para elección aleatoria.

4.2. Correlación promedio vs Valor propio más grande

A continuación se estudiará la relación entre el valor de correlación promedio de la matriz C tanto con el valor propio más grande de C como con el valor propio más grande de los dos ensembles estudiados.

4.2.1. $\langle C_{ij} \rangle$ vs λ_1

Antes de pasar a lo que se obtiene con el modelo en sí, se verá como está relacionada la correlación promedio $\langle C_{ij} \rangle$ con el valor propio más grande de la matriz de correlación C . La correlación promedio $\langle C_{ij} \rangle$ se define como

$$\langle C_{ij} \rangle = \frac{1}{N(N-1)} \left(\sum_{ij} C_{ij} - N \right), \quad (4.4)$$

ya que solo se toman en cuenta los elementos no diagonales de C .

En la Fig. 4.20 se puede ver como evoluciona la correlación promedio $\langle C_{ij} \rangle$ (izquierda) y el valor propio más grande λ_1 (derecha) de la matriz de correlación completa. A primera vista parece que se puede aplicar el modelo de correlación constante, del que se hablo en el la sección de Ensemble de Wishart Correlacionado, pero como se muestra en la Fig. 4.21, al calcular la correlación estimada C_{est} a partir de la Ec. 4.1 no se obtiene una coincidencia perfecta, aunque cercana.

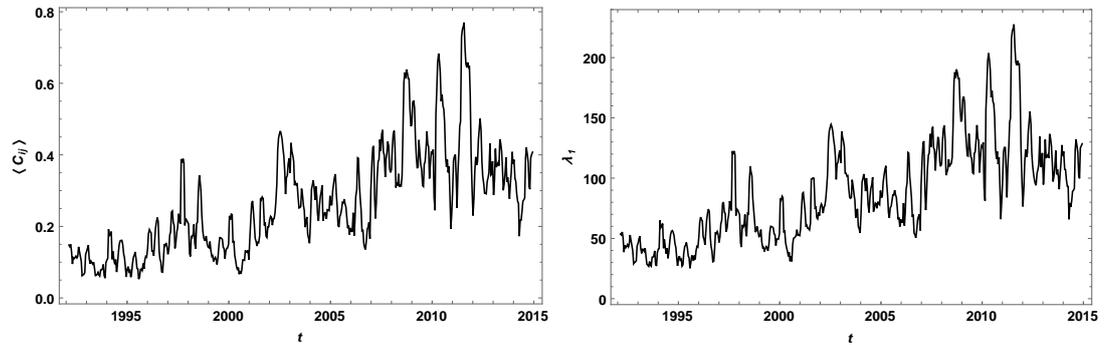


Figura 4.20: Evolución de la correlación promedio (izquierda) y el valor propio más grande (derecha) de C .

A pesar de no tratarse de un modelo de correlación constante, claramente hay una relación lineal entre $\langle C_{ij} \rangle$ y λ_1 como se muestra en la Fig. 4.22. Por lo que se tiene,

$$\lambda_1 = (281,98 \pm 0,73) \langle C_{ij} \rangle + (12,69 \pm 0,22), \quad (4.5)$$

4.2. CORRELACIÓN PROMEDIO VS VALOR PROPIO MÁS GRANDE³³

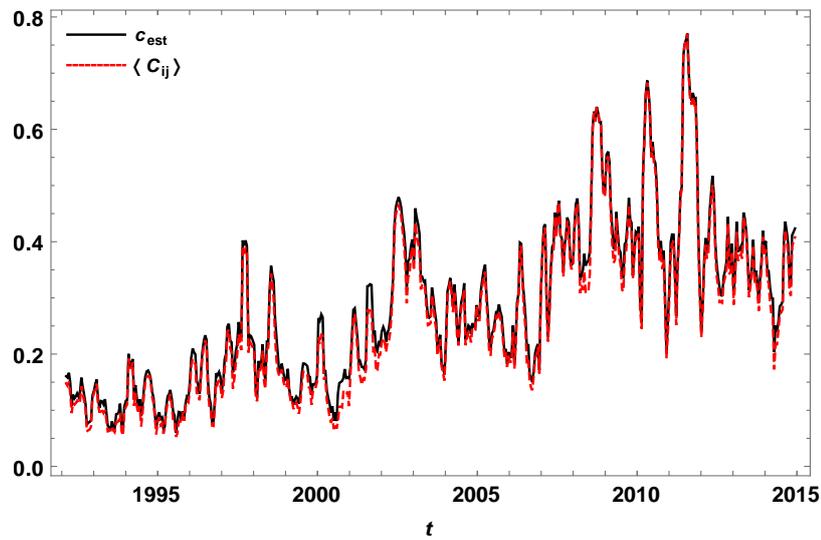


Figura 4.21: Comparación entre la correlación promedio (curva continua) y la correlación estimada (curva punteada) tomando el modelo de correlación constante.

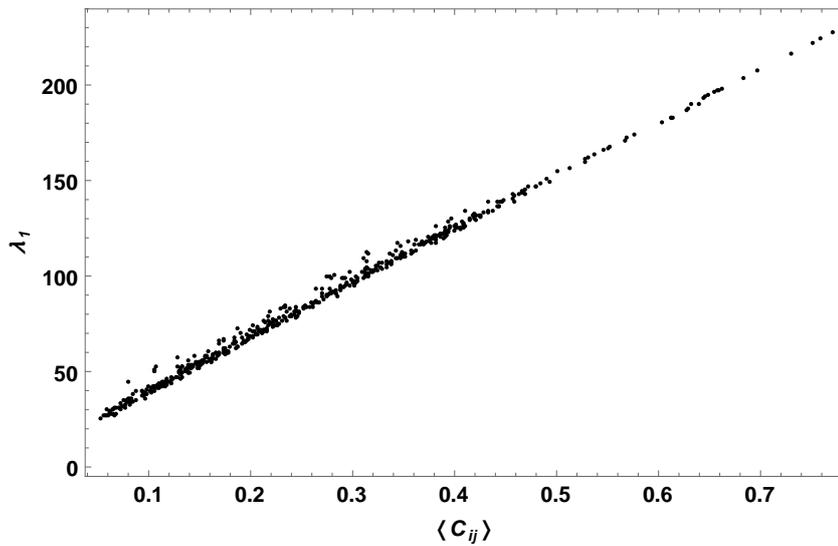


Figura 4.22: Relación lineal entre $\langle C_{ij} \rangle$ y λ_1 .

en nuestro caso para $N = 293$ y $T = 44$. Aunque anteriormente se ha visto que el valor más grande destaca del resto [5], y representa al mercado, no se ha reportado que la relación entre éste y la correlación promedio sea lineal a lo largo del tiempo.

4.2.2. $\langle C_{ij} \rangle$ vs $\langle \lambda_1 \rangle_{R,S}$

Tanto $\langle \lambda_1 \rangle_R$ como $\langle \lambda_1 \rangle_S$ se comportan como la correlación promedio cuando se trabaja con correlaciones ordinarias. Y como se muestra en la Fig. 4.23, ambos guardan una relación lineal con la correlación promedio. Las

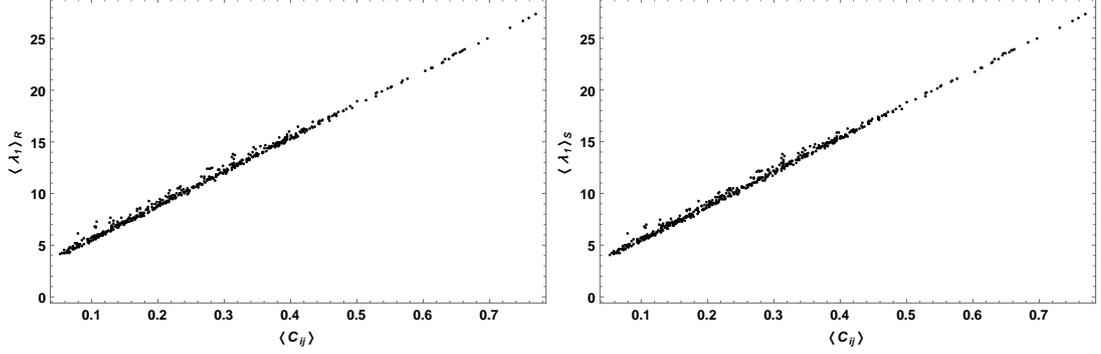


Figura 4.23: Relación lineal entre $\langle C_{ij} \rangle$ y $\langle \lambda_1 \rangle_R$ (izquierda) y $\langle \lambda_1 \rangle_S$ (derecha).

relaciones lineales son,

$$\langle \lambda_1 \rangle_R = (32,54 \pm 0,07) \langle C_{ij} \rangle + (2,43 \pm 0,02)$$

$$\langle \lambda_1 \rangle_S = (32,65 \pm 0,07) \langle C_{ij} \rangle + (2,35 \pm 0,02), \quad (4.6)$$

para $n = 35$ y $T = 44$.

Pero cuando se usa como referencia a la matriz de correlaciones parciales, el valor propio más grande no tiene relación lineal con el valor de correlación promedio de C ni con la correlación promedio de la matriz de correlaciones parciales.

4.2.3. Relación entre λ_1 y $\langle \lambda_1 \rangle_{R,S}$

En la Fig. 4.24 se puede observar que el valor propio λ_1 de la matriz C tiene una relación lineal muy clara con el valor promedio de los valores propios más grandes tanto del ensemble aleatorio $\langle \lambda_1 \rangle_R$ como del ensemble por sectores $\langle \lambda_1 \rangle_S$. Dicha relación es:

$$\lambda_1 = (8,666 \pm 0,004) \langle \lambda_1 \rangle_R - (8,42 \pm 0,05)$$

$$\lambda_1 = (8,636 \pm 0,006) \langle \lambda_1 \rangle_S - (7,63 \pm 0,08). \quad (4.7)$$

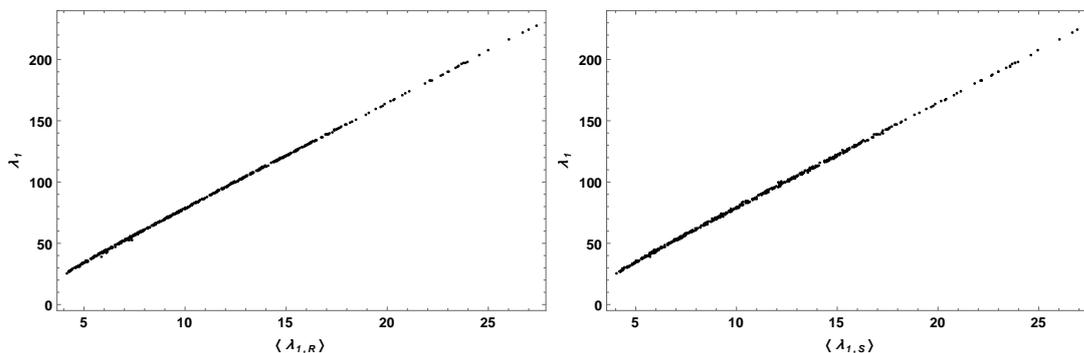


Figura 4.24: Relación lineal entre λ_1 y $\langle \lambda_1 \rangle_R$ (izquierda) y $\langle \lambda_1 \rangle_S$ (derecha).

El que obtengamos una relación lineal perfecta, entre el máximo valor propio de la matriz C y el valor promedio de la distribución para el máximo valor propio de los ensembles calculados, quiere decir que en promedio las matrices de correlación C_α preservan la estructura y las características de C .

4.3. Distribución de espaciamiento a primer vecino

Una de las principales razones del uso del método propuesto es que se obtiene suficientes valores propios para hacer un análisis estadístico. Donde realmente se ve la ventaja del método es en el cálculo de funciones tales como la distribución del espaciamiento a primer vecino $p(s)$ y la varianza de número $\Sigma^2(L)$, y como evolucionan en el tiempo.

En general, la distribución de espaciamiento a primer vecino $p(s)$ se ve como en la Fig. 4.25, y es difícil distinguir a ojo cuales son los cambios en la distribución entre un tiempo y otro. Esta observación es válida tanto para correlaciones ordinarias como para correlaciones parciales. Una forma de ver como cambia la distribución $p(s)$ es calculando la desviación estándar σ de la distribución, por ejemplo. La forma en la que evoluciona σ para ambos tipos de correlación se muestra en la Fig. 4.26.

Algo que se observa es que el efecto de situación económica del mercado está reflejado en la distribución $P(s)$, ya que al parecer en momentos de crisis, como en el 2001-2002 y 2008, la desviación estándar aumenta. Que σ aumente significa que las fluctuaciones en el espectro a corto alcance se hacen más grandes.

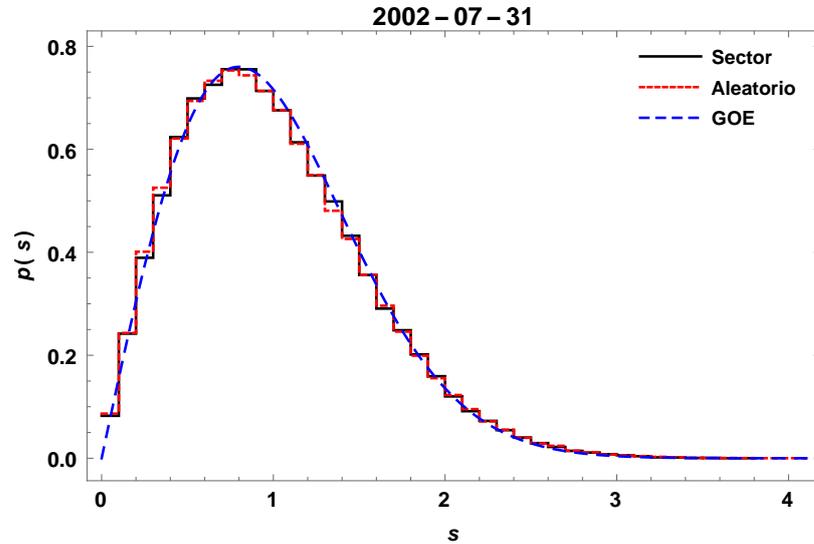


Figura 4.25: Distribución $p(s)$ para $t = 1993 - 05 - 25$. Para correlaciones ordinarias. Las curvas escalonadas son para los ensembles y la curva rallada es la distribución para la conjetura de Wigner para el GOE dada por Ec. 2.9.

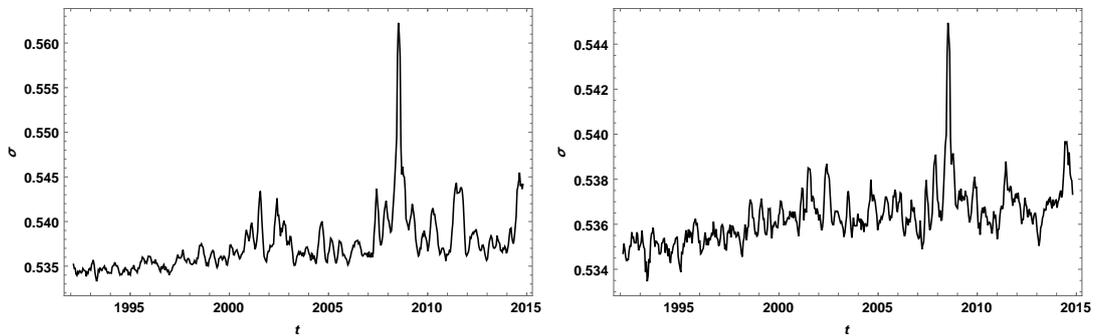


Figura 4.26: Comparación entre la evolución de la desviación estándar σ de la distribución $p(s)$ cuando se utiliza correlación ordinaria (izquierda) y parcial (derecha). Se muestran resultados para el ensemble por sector, dado que el otro caso es muy parecido.

4.4. Varianza de número

La varianza de número $\Sigma^2(L)$, al igual que las otras distribuciones antes mencionadas, cambia en el tiempo, y aunque en distintos periodos de tiempo se parece a la obtenida para el GOE, en otros se desvía bastante.

Antes que nada hay que decir que compararemos con la varianza de número obtenida para un ensemble GOE de elementos con dimensión 35×35 , al que

llamaremos GOE Est. La diferencia entre la formula analítica para $\Sigma^2(L)$ del ensemble GOE (Ec. 2.13) y la varianza de número para GOE Est se muestra en la Fig. 4.27. Estas curvas coinciden para valores pequeños de L pero no para valores grandes, pues GOE Est cae más rápido.

Algunos ejemplos de como es la varianza de número para distintos tiempos se presentan en las figuras 4.28, 4.29 y 4.30. En general $\Sigma^2(L)$ se acerca más al GOE Est cuando se trata con correlaciones parciales. En la Fig. 4.28 se

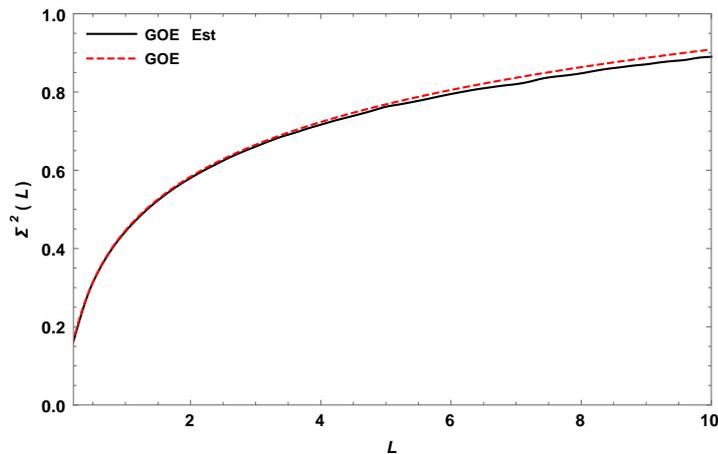


Figura 4.27: Comparación entre la forma analítica de $\Sigma^2(L)$ para el GOE y la varianza de número obtenida numéricamente para matrices de dimensión 35×35 , GOE Est.

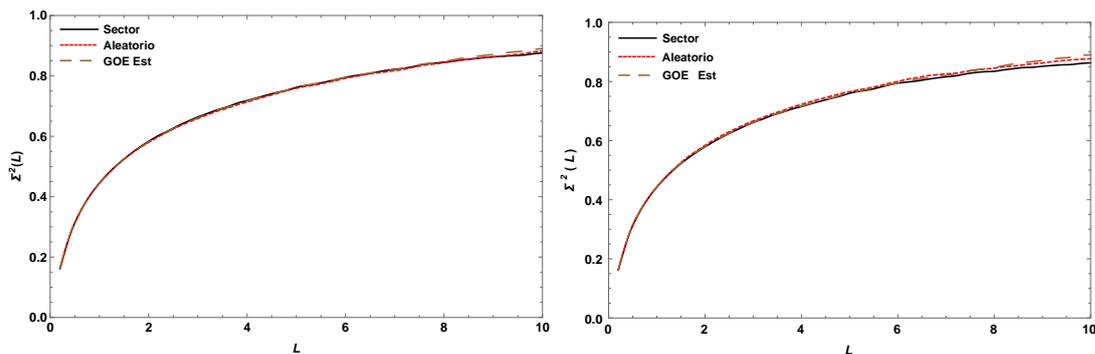


Figura 4.28: Varianza de número para $t = 1993 - 05 - 25$. Cuando se usan correlaciones ordinarias (izquierda) y cuando se usa correlaciones parciales (derecha).

observa que la varianza de número tanto para correlaciones ordinarias como para correlaciones parciales se acercan a GOE Est, además esta fecha tiene un valor pequeño para el valor propio más grande (Fig. 4.6).

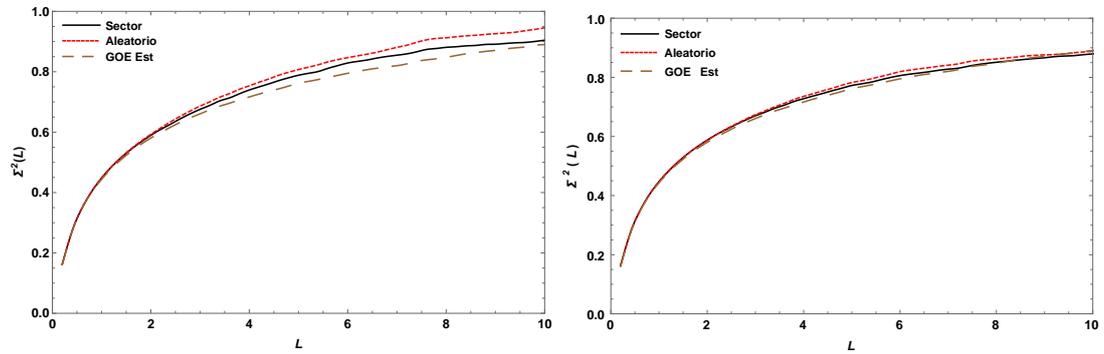


Figura 4.29: Varianza de número para $t = 2002 - 07 - 31$. Cuando se usan correlaciones ordinarias (izquierda) y cuando se usa correlaciones parciales (derecha).

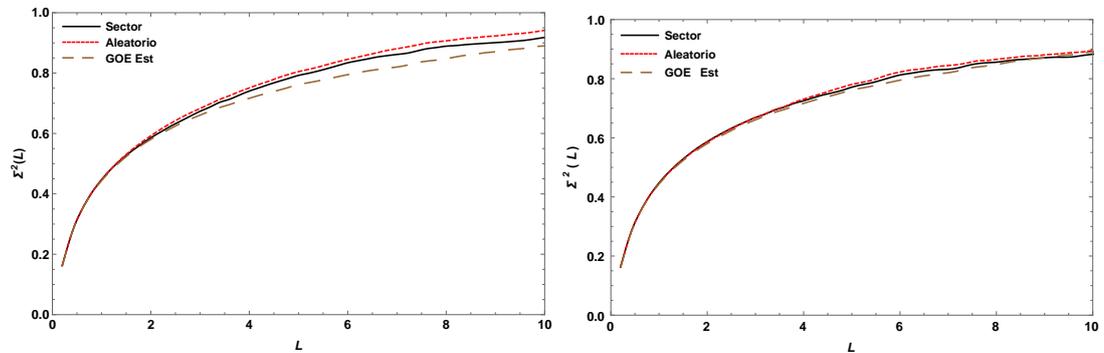


Figura 4.30: Varianza de número para $t = 2009 - 07 - 13$. Cuando se usan correlaciones ordinarias (izquierda) y cuando se usa correlaciones parciales (derecha).

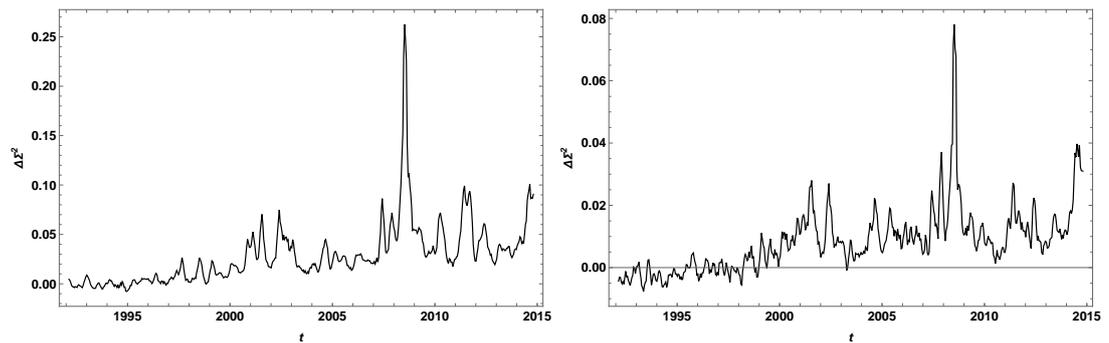


Figura 4.31: Comparación entre la evolución de la diferencia 4.8 con correlación ordinaria (izquierda) y parcial (derecha). Para el ensemble por sector, y el otro caso es muy parecido.

Para medir cuánto se separa del caso GOE Est se calcula la integral de

la diferencia de las curvas, de la siguiente forma:

$$\Delta\Sigma^2 = \left(\int \Sigma^2(L)dL - \int \Sigma_{\text{GOE Est}}^2(L)dL \right) / \int \Sigma_{\text{GOE Est}}^2(L)dL. \quad (4.8)$$

Esta diferencia se muestra en la Fig. 4.31, y se puede ver que se tiene un comportamiento similar tanto para el caso con correlación ordinaria como con correlación parcial. Además, estas curvas son similares a las que se obtuvieron para la desviación estándar σ de la distribución $p(s)$.

Cabe resaltar que aunque el método con correlaciones ordinarias y con correlaciones parciales difiere claramente para la evolución del valor propio más grande, esto no es así para las distribuciones $p(s)$ y $\Sigma^2(L)$, ya que muestran un comportamiento cualitativamente similar.

Capítulo 5

Conclusiones

Al lo largo del trabajo se estudió un nuevo método para analizar series de tiempo cortas que permite obtener más valores propios, con los que se puede hacer un mejor análisis estadístico de estos. El método se usó tanto para caracterizar correlaciones ordinarias como correlaciones parciales, en donde se remueve el efecto de los rendimientos promedio.

Se observó que los valores propios se separan principalmente en dos distribuciones para el caso de correlación ordinaria: el bulk, que está formado por todos los valores propios salvo el más grande y donde una distribución para el ensemble de Wishart correlacionado es más apropiada; y la distribución para el valor propio más grande, que se separa del resto conforme la correlación crece. Se clasificó así porque el valor propio más grande se separa significativamente.

A pesar de esta clasificación, sí hay otros valores propios que se separan del resto, aunque no es tan notorio como en el caso del mayor valor propio. Que se alejen otros valores propios o no, depende de la estructura de la matriz de correlación.

Se determinó que existe una relación lineal clara entre el valor promedio de la distribución del valor propio más grande $\langle \lambda_1 \rangle_{R,S}$ y la correlación promedio $\langle C_{ij} \rangle$. De la misma forma, hay una relación lineal entre el máximo valor propio λ_1 de la matriz C y $\langle \lambda_1 \rangle_{R,S}$.

A pesar de que el comportamiento del valor propio más grande difiere dependiendo de si se usan correlaciones ordinarias o correlaciones parciales, se observó que las fluctuaciones espectrales, tales como la distribución de espaciamiento a primer vecino y la varianza de número, se comportan de manera similar. Además de que parece haber una relación entre la desviación estándar de la distribución $p(s)$ y la diferencia de la varianza de número del mercado con la varianza de número del GOE (Est).

Por otro lado, no se observaron diferencias sobresalientes entre la construcción del ensemble de forma aleatoria y tomando en cuenta los sectores industriales.

Apéndice A

Desdoblamiento

El desdoblamiento, o mejor conocido como *unfolding*, es una técnica que permite analizar las fluctuaciones de un espectro [9].

Supón que el resultado de una serie de mediciones arroja una secuencia de valores $\{E_1, E_2, \dots, E_N\}$, que puede ser representada por la función espectral

$$S(E) = \sum_{n=1}^N \delta(E - E_n). \quad (\text{A.1})$$

Para estudiar las propiedades de las fluctuaciones del espectro se tiene que remover la densidad central de niveles $R_1(E)$, a este procedimiento se le conoce como desdoblamiento. Se define la función espectral comulativa como

$$\eta(E) = \int_{-\infty}^E S(E') dE' = \sum_{n=1}^N \Theta(E - E_n). \quad (\text{A.2})$$

Esta función cuenta el número de niveles menores o iguales a E . Esta función es separada en una parte suave $\xi(E)$ y otra con fluctuaciones $\eta_{fl}(E)$,

$$\eta(E) = \xi(E) + \eta_{fl}(E). \quad (\text{A.3})$$

La parte suave esta dada por

$$\xi(E) = \int_{-\infty}^E R_1(E') dE'. \quad (\text{A.4})$$

Para desdoblar el espectro, la secuencia $\{E_1, E_2, \dots, E_N\}$ se mapea a los números $\{\xi_1, \xi_2, \dots, \xi_N\}$ con

$$\xi_n = \xi(E_n), \quad n = 1, \dots, N. \quad (\text{A.5})$$

De esta forma la función cumulativa para las nuevas variables es

$$\hat{\xi}(E) = \xi + \hat{\eta}_{fl}(\xi). \quad (\text{A.6})$$

Por lo tanto, la densidad central de niveles de las variables nuevas es la unidad, ya que es la derivada de la parte suave.

Además, la función de correlación R_k tiene que ser reescalada apropiadamente. La función de correlación desdoblada X_k se obtiene al igual las probabilidades diferenciales en ambas escalas

$$X_k(\xi_1, \dots, \xi_k) d\xi_1 \cdots d\xi_k = R_k(\lambda_1, \dots, \lambda_k) d\lambda_1 \cdots d\lambda_k.$$

En la práctica, desdoblar un espectro es algo técnico. Se ajusta un polinomio a la función cumulativa que capture sus principales características, y se evalúa el espectro en el polinomio. Dando como resultado las nuevas variables.

Bibliografía

- [1] R. Cont, “Empirical properties of asset returns: stylized facts and statistical issues,” *Quantitative Finance*, vol. 1, pp. 223–236, 2001.
- [2] M. C. Münnix, T. Shimada, R. Schäfer, F. Leyvraz, T. H. Seligman, T. Guhr, and H. E. Stanley, “Identifying states of a financial market,” *Scientific Reports*, vol. 2, p. 644, Sep 2012.
- [3] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, T. Guhr, and H. E. Stanley, “Random matrix approach to cross correlations in financial data,” *Phys. Rev. E*, vol. 65, p. 066126, Jun 2002.
- [4] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. Nunes Amaral, and H. E. Stanley, “Universal and nonuniversal properties of cross correlations in financial time series,” *Physical Review Letters*, vol. 83, pp. 1471–1474, Aug 1999.
- [5] L. Laloux, P. Cizeau, J.-P. Bouchaud, and M. Potters, “Noise dressing of financial correlation matrices,” *Phys. Rev. Lett.*, vol. 83, pp. 1467–1470, Aug 1999.
- [6] M. Potters, J.-P. Bouchaud, and L. Laloux, “Financial applications of random matrix theory: Old laces and new pieces,” *Acta Physica Polonica B*, vol. 36, p. 2767, Sep 2005.
- [7] Vinayak, R. Schäfer, and T. H. Seligman, “Emerging spectra of singular correlation matrices under small power-map deformations,” *Phys. Rev. E*, vol. 88, p. 032115, sep 2013.
- [8] T. A. Brody, J. Flores, J. B. French, P. A. Mello, A. Pandey, and S. S. M. Wong, “Random-matrix physics: spectrum and strength fluctuations,” *Rev. Mod. Phys.*, vol. 53, pp. 385–479, Jul 1981.
- [9] T. Guhr, A. M. Groeling, and H. A. Weidenmueller, “Random-matrix theories in quantum physics: common concepts,” *Physics Reports*, vol. 299, no. 4, pp. 189 – 425, 1998.

- [10] J. Wishart, “The generalised product moment distribution in samples from a normal multivariate population,” *Biometrika*, vol. 20A, no. 1/2, pp. 32–52, 1928.
- [11] V. A. Marčenko and L. A. Pastur, “Distribution of eigenvalues for some sets of random matrices,” *Mathematics of the USSR-Sbornik*, vol. 1, no. 4, p. 457, 1967.
- [12] G. Enderlein, “Wilks, s. s.: Mathematical statistics. j. wiley and sons, new york–london 1962; 644 s., 98 s,” *Biometrische Zeitschrift*, vol. 6, no. 3, pp. 214–215, 1964.
- [13] R. J. Muirhead, *Aspects of multivariate statistical theory*. Wiley series in probability and mathematical statistics. Probability and mathematical statistics, New York: John Wiley and Sons, 2008.
- [14] Vinayak and A. Pandey, “Correlated wishart ensembles and chaotic time series,” *Phys. Rev. E*, vol. 81, p. 036202, Mar 2010.
- [15] K. Baba, R. Shibata, and M. Sibuya, “Partial correlation and conditional correlation as measures of conditional independence,” *Australian and New Zealand Journal of Statistics*, vol. 46, no. 4, pp. 657–664, 2004.
- [16] E. P. Wigner, “Characteristic vectors of bordered matrices with infinite dimensions,” *Annals of Mathematics*, vol. 62, no. 3, pp. 548–564, 1955.
- [17] E. P. Wigner, “Characteristics vectors of bordered matrices with infinite dimensions ii,” *Annals of Mathematics*, vol. 65, no. 2, pp. 203–207, 1957.
- [18] E. P. Wigner, “On the distribution of the roots of certain symmetric matrices,” *Annals of Mathematics*, vol. 67, no. 2, pp. 325–327, 1958.
- [19] A. T. James, “Normal multivariate analysis and the orthogonal group,” *The Annals of Mathematical Statistics*, vol. 25, no. 1, pp. 40–75, 1954.
- [20] M. L. Mehta, *Random Matrices; 3rd ed.* Pure and applied mathematics series, San Diego, CA: Elsevier, 2004.
- [21] S. Ghosh and A. Pandey, “Skew-orthogonal polynomials and random-matrix ensembles,” *Phys. Rev. E*, vol. 65, p. 046221, Apr 2002.
- [22] A. Pandey and S. Ghosh, “Skew-orthogonal polynomials and universality of energy-level correlations,” *Phys. Rev. Lett.*, vol. 87, p. 024102, Jun 2001.

- [23] C. Recher, M. Kieburg, and T. Guhr, “Eigenvalue Densities of Real and Complex Wishart Correlation Matrices,” *Physical Review Letters*, vol. 105, p. 244101, Dec. 2010.
- [24] C. Recher, M. Kieburg, T. Guhr, and M. R. Zirnbauer, “Supersymmetry Approach to Wishart Correlation Matrices: Exact Results,” *Journal of Statistical Physics*, vol. 148, pp. 981–998, Sept. 2012.
- [25] D. Waltner, T. Wirtz, and T. Guhr, “Eigenvalue density of the doubly correlated Wishart model: exact results,” *Journal of Physics A Mathematical General*, vol. 48, p. 175204, may 2015.
- [26] A. Bohr and B. R. Mottelson, *Nuclear Structure. Single-Particle Motion.*, vol. 1. Benjamin, New York, 1969.
- [27] J. Campbell, A. Lo, and A. MacKinlay, *The Econometrics of Financial Markets*. Princeton University Press, 1997.
- [28] G. W. Schwert, “Stock volatility during the recent financial crisis,” Working Paper 16976, National Bureau of Economic Research, April 2011.
- [29] R. Shiller, *Market Volatility*. MIT Press, 1992.
- [30] W.-L. Lin, R. F. Engle, and T. Ito, “Do bulls and bears move across borders? international transmission of stock returns and volatility as the world turns,” Working Paper 3911, National Bureau of Economic Research, November 1991.
- [31] Y. L. F. Bruno Solnik, Cyril Boucrelle, “International market correlation and volatility,” *Financial Analysts Journal*, vol. 52, no. 5, pp. 17–34, 1996.
- [32] C. B. Erb, C. R. Harvey, and T. E. Viskanta, “Forecasting international equity correlations,” *Financial Analysts Journal*, vol. 50, no. 6, p. 32, 1994.
- [33] S. Drożdż, F. Grümmer, A. Górski, F. Ruf, and J. Speth, “Dynamics of competition between collectivity and noise in the stock market,” *Physica A: Statistical Mechanics and its Applications*, vol. 287, no. 3–4, pp. 440 – 449, 2000.
- [34] L. S. Junior and I. D. P. Franca, “Correlation of financial markets in times of crisis,” Papers 1102.1339, arXiv.org, Feb. 2011.

- [35] S. M. Bartram and Y.-H. Wang, “Another look at the relationship between cross-market correlation and volatility,” *Finance Research Letters*, vol. 2, no. 2, pp. 75 – 88, 2005.
- [36] L. Borland and Y. Hassid, “Market panic on different time-scales,” *ArXiv e-prints*, Oct 2010.
- [37] G. W. Schwert, “Why does stock market volatility change over time?,” *The Journal of Finance*, vol. 44, no. 5, pp. 1115–1153, 1989.
- [38] B. Mandelbrot, “The variation of certain speculative prices,” *The Journal of Business*, vol. 36, no. 4, pp. 394–419, 1963.
- [39] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*. Communications and signal processing, McGraw-Hill, 1991.
- [40] R. Mantegna and H. Stanley, *Introduction to Econophysics: Correlations and Complexity in Finance*. Cambridge University Press, 1999.
- [41] A. De La Fuente, N. Bing, I. Hoeschele, and P. Mendes, “Discovery of meaningful associations in genomic data using partial correlation coefficients,” *Bioinformatics*, vol. 20, no. 18, pp. 3565–3574, 2004.
- [42] A. Reverter and E. K. Chan, “Combining partial correlation and an information theory approach to the reversed engineering of gene co-expression networks,” *Bioinformatics*, vol. 24, no. 21, pp. 2491–2497, 2008.
- [43] L. Chen, S. Zheng, *et al.*, “Studying alternative splicing regulatory networks through partial correlation analysis,” *Genome Biol*, vol. 10, no. 1, p. R3, 2009.
- [44] Y. Shapira, D. Y. Kenett, and E. Ben-Jacob, “The index cohesive effect on stock market correlations,” *The European Physical Journal B*, vol. 72, no. 4, pp. 657–669, 2009.
- [45] H. Meng, W.-J. Xie, Z.-Q. Jiang, B. Podobnik, W.-X. Zhou, and H. E. Stanley, “Systemic risk and spatiotemporal dynamics of the us housing market,” *Scientific reports*, vol. 4, 2014.
- [46] D. Y. Kenett, X. Huang, I. Vodenska, S. Havlin, and H. E. Stanley, “Partial correlation analysis: applications for financial markets,” *Quantitative Finance*, vol. 15, no. 4, pp. 569–578, 2015.