



**TECNOLÓGICO
DE MONTERREY**

Instituto Tecnológico y de Estudios
Superiores de Monterrey
Campus Ciudad de México

*Estudio de la distribución hiperbólica generalizada
multivariada: aspectos teóricos y numéricos para
aplicaciones financieras*

TESIS QUE PARA RECIBIR EL TÍTULO DE
DOCTORADO EN CIENCIAS FINANCIERAS
PRESENTA

Leovardo Mata Mata



**TECNOLÓGICO
DE MONTERREY**

Director de tesis:
José Antonio Núñez Mora

Biblioteca
Campus Ciudad de México

Lectores: José Luis de la Cruz Gallegos y José Jorge Mora Rivera

México D.F., 08 de Noviembre de 2013

RESUMEN

La matriz de covarianza empírica para un conjunto de rendimientos en una serie de activos es una herramienta de gran importancia en el área de las finanzas, tanto para la administración de riesgos como para la administración de portafolios. Sin embargo, en la práctica resulta difícil determinar una matriz de covarianza empírica confiable, ya que la estimación está contaminada por el “ruido” de medición presente en las series de tiempo recopiladas. En esta tesis se abordan diferentes posibilidades para distinguir la “señal de ruido”, y así filtrar la matriz de covarianzas que contiene información verdadera para el control de riesgos.

Una alternativa para analizar la variación del ruido y sus efectos en la matriz de covarianzas consiste en una colección de técnicas empleadas en física estadística. Estos procedimientos no paramétricos se engloban bajo la teoría de matrices aleatorias y permiten, bajo ciertas pruebas de hipótesis, discernir la variación por ruido. No obstante, no añade información sobre la distribución de probabilidad conjunta subyacente, por ello se comparan sus resultados con la estimación paramétrica de una distribución de probabilidad altamente versátil, la familia de distribuciones hiperbólicas multivariadas generalizadas. En ese caso, la matriz de covarianzas empírica puede obtenerse como un estimador que bajo los parámetros de la distribución de probabilidad captura algunos de los hechos estilizados más sobresalientes de los rendimientos financieros, tales como: colas pesadas, asimetría y alta curtosis; sin mencionar la no normalidad.

Sin embargo, uno de los obstáculos para la aplicación de estas distribuciones multivariantes hiperbólicas generalizadas es la dificultad numérica para estimar y calibrar los parámetros de la distribución sobre un conjunto de datos. Por ello, realizamos una revisión de los algoritmos que permiten ajustar las distribuciones GH, pero bajo un espectro más amplio de parámetros de lo que usualmente se cita en la literatura. En particular, el parámetro de la función de Bessel de tercer orden no se asume fijo, y el procedimiento numérico comienza con un punto inicial elegido sistemáticamente por las propiedades de invarianza de las combinaciones lineales de la familia GH. Esta es una modificación de los métodos propuestos en McNeil, Frey, y Embrechts (2005), Protasov (2004), Hu (2005), y Lüthi y Wolfgang (2011).

El procedimiento de estimación se complementa con una prueba de hipótesis empírica para distribuciones de probabilidad continuas multivariadas, McAssey (2013). De esta manera es posible valorar la bondad de ajuste de los parámetros calculados. Estos algoritmos modificados extienden conjuntamente la estabilidad de los procedimientos para ajustar las funciones de distribución de probabilidad y facilitan su aplicación a diversas áreas de la optimización de portafolios y la administración de riesgos.

Asimismo, se presentan aplicaciones de los algoritmos de estimación modificados en diferentes contextos. Primero, se contrastan los resultados presentados en Protasov (2004) para un conjunto de cinco tipos de cambio, frente a las estimaciones con un parámetro libre de la función de Bessel de tercer orden. En segundo lugar, se presenta la estimación de un portafolio de mínima varianza mediante la distribución de probabilidad multivariada hiperbólica generalizada para una muestra de acciones de NASDAQ, NYSE y algunos commodities, en contraste con las estimaciones usuales de Markowitz y una matriz de correlaciones filtrada por matrices aleatorias. En tercer lugar se estima la distribución para dos diferentes periodos y se muestra como la variación del rendimiento de los commodities metálicos resulta de mayor relevancia después de la crisis de finales de la década anterior.

ÍNDICE GENERAL

Capítulo 1. Introducción	1
Capítulo 2. Enfoques estadísticos	9
2.1. Matrices aleatorias	9
2.2. La distribución hiperbólica generalizada	16
2.2.1. La función de Bessel de tercer orden	16
2.2.2. Distribución inversa gaussiana generalizada	18
2.2.3. Distribución hiperbólica generalizada univariada	21
2.2.4. Función de distribución hiperbólica generalizada multivariada	25
Capítulo 3. Estimación de la distribución hiperbólica generalizada multivariada	29
3.1. El algoritmo de estimación EM	30
3.2. El algoritmo EM modificado	37
3.3. Una prueba empírica para distribuciones multivariadas continuas	40
Capítulo 4. Aplicaciones y ejemplos	45
4.1. Aspectos numéricos para estimar la distribución GH	45
4.2. Variación del rendimiento de diferentes commodities durante Periodo 2000-2012	49
4.3. Herramientas complementarias para construir portafolios eficientes	59
Conclusiones	69
Referencias	71
Anexos	76

LISTA DE FIGURAS

Figura 2.1	11
Figura 2.2	19
Figura 2.3	22
Figura 4.1	52
Figura 4.2	54
Figura 4.3	57
Figura 4.4	58
Figura 4.5	68

LISTA DE CUADROS

Cuadro 2.1	21
Cuadro 2.2	23
Cuadro 4.1	47
Cuadro 4.2	48
Cuadro 4.3.....	55
Cuadro 4.4.....	56
Cuadro 4.5.....	56
Cuadro 4.6.....	65
Cuadro 4.7.....	66
Cuadro 4.8.....	67
Cuadro 4.9.....	67

Capítulo 1

Introducción

El estudio de la covarianza entre diferentes variables tiene una larga historia en las finanzas, ya que la matriz de covarianzas es el insumo fundamental de la teoría de optimización de portafolios, Markowitz (1959). Sabemos que medir la covarianza entre diferentes rendimientos es relevante para tener una medida de la asociación que existe entre las variables de un cierto portafolio, pues la posibilidad de grandes pérdidas para una cierta cartera está determinada por los movimientos correlacionados de sus diferentes elementos.

Específicamente, el rendimiento promedio de un portafolio puede pensarse como una combinación lineal de los rendimientos individuales de los N activos involucrados, siendo que el monto de capital invertido en cada activo debe determinarse de tal manera que el riesgo, medido por la varianza total, sea lo más pequeño posible dado un nivel de rendimiento dado. Este portafolio óptimo, puede fácilmente encontrarse mediante multiplicadores de Lagrange, Engels (2004). En ese caso se puede observar que la solución depende de la matriz inversa de covarianzas y por tanto, la estimación de la matriz de covarianzas resulta crucial para hallar los montos de capital adecuados en el portafolio óptimo. No obstante, en la práctica resulta difícil determinar una matriz de covarianza

empírica confiable, Laloux, Cizeau, Potters y Bouchaud (2000).

La estimación de una matriz de covarianzas para un conjunto de N diferentes activos requiere estimar $N(N - 1) / 2$ parámetros, los cuales deben determinarse a partir de N series de tiempo de longitud T . Si T no es muy grande en comparación con N , se puede esperar que la determinación de las covarianzas sea “ruidosa”, y por lo tanto que la matriz de covarianzas empírica sea en gran medida aleatoria, pues no refleja información sustancialmente verdadera, Burda (2004). En esa situación, la estructura de la matriz de covarianzas está dominada por el ruido de medición presente en las series de tiempo recopiladas, así que debemos ser muy cuidadosos al utilizar la matriz de covarianzas en aplicaciones. Esto nos dice que resulta relevante idear métodos que permitan distinguir la "señal" de ruido”, para así filtrar la matriz de covarianzas que contiene información verdadera e incluirla para el control de riesgos. De esa forma descartamos la información no útil y que resulta inestable en el tiempo, Laloux et al (1999).

Sabemos que el portafolio óptimo construido bajo el enfoque de Markowitz emplea la matriz inversa de covarianzas, esto implica por las propiedades matriciales de la matriz de covarianzas, Engels (2004), que la solución para un portafolio óptimo depende de sus valores y vectores propios. En particular, podemos percatarnos que son los valores propios más pequeños quienes son más sensibles al ruido, y por ende sus correspondientes vectores propios afectan directamente a la cartera óptima de menor riesgo.

Existen técnicas estadísticas que incluyen en sus procedimientos valores y vectores propios para analizar la varianza total, a saber, análisis de componentes principales (ACP) y análisis factorial (AF). Algunos autores consideran ACP como una etapa de AF y otros consideran que se tratan de técnicas distintas, Jhonson (2007). ACP trata de hallar componentes (factores) que sucesivamente expliquen la mayor parte de la varianza total. En contraste, AF busca factores que expliquen la mayor parte de la varianza total

común. AF distingue entre varianza común y varianza única, mientras que ACP no hace distinción y se centra en la varianza, Jhonson (2007).

El procedimiento ACP busca hallar combinaciones lineales de las variables originales que expliquen la mayor parte de la variación total, mientras que AF pretende encontrar un nuevo conjunto de variables, menor en número que las variables originales y que expresen aquello que es común a dichas variables. Existen procedimientos numéricos extendidos de ACP y AF mediante simulación para discernir que componentes (factores) capturan información genuina libre de ruido de las variables originales, Bruin (2006), pero se centran en los valores propios de magnitud mayor y no caracterizan que sucede con las variaciones expresadas por los valores propios más pequeños.

Si deseamos refinar en mayor medida la estimación de una matriz de covarianzas, requerimos estudiar la variación total más allá de ACP o AF, de tal forma que utilicemos todos los valores propios calculados. En otras palabras, buscamos analizar las propiedades de la matriz de covarianzas empírica bajo la hipótesis nula de una matriz de covarianzas aleatoria. El objetivo es distinguir en la matriz de covarianzas empírica la variación de los rendimientos atribuible a “ruido” y la parte que contiene efectivamente “información”, Burda (2004).

En ese contexto, sabemos que en física estadística existen un conjunto de técnicas asociadas con teoría de matrices aleatorias (TMA) para discernir el ruido presente en un conjunto de variables. Si bien es cierto que no describen la distribución de probabilidad conjunta subyacente a todas las variables involucradas, sí permiten realizar un conjunto de pruebas de hipótesis sobre los valores propios para distinguir la magnitud del ruido en una muestra aleatoria y filtrar la matriz de covarianzas empíricas, Brody (1981).

El uso y estudio de las matrices aleatorias para clasificar la variación por ruido en una muestra aleatoria se remonta a la física nuclear, donde se buscaba explicar estadís-

ticamente el comportamiento de los niveles de energía en sistemas cuánticos complejos, Mehta (2004). Más recientemente, se ha mostrado que estos resultados tienen aplicaciones importantes en el contexto financiero, Maslow (2001) y Laloux et al. (1999).

Laloux et al. (2000) emplea las pruebas de hipótesis y los procedimientos de la TMA para analizar las correlaciones entre diferentes stocks del índice Standar and Poors 500 (S&P500), y justifica que las propiedades de la TMA son aplicables en las matrices de correlación empíricas. En ese sentido, podemos aplicar la TMA para discernir la variación de los rendimientos atribuible a ruido para filtrar la matriz de covarianzas y refinar la estimación óptima de un portafolio de Markowitz.

Los procedimientos de matrices aleatorias permiten refinar la estimación de la matriz de covarianzas bajo pruebas de hipótesis que satisfacen criterios asintóticos más generales que aquellos asociados con ACP o AF, Maslow (2001). TMA no sólo permite distinguir la mayor variación común en un conjunto de variables mediante los valores propios, sino que describe que ocurre con los valores propios más pequeños, lo cual no es posible con ACP o AF. No obstante, TMA no permite ahondar en las relaciones existentes entre las diferentes variables, tales como colas pesadas, alta curtosis, sesgo, entre otras. Para ello, sería deseable estimar la distribución de probabilidad multivariada implícita y emplear la matriz de covarianzas poblacional determinada por la distribución de probabilidad subyacente al conjunto de variables. Sin mencionar que para garantizar confiabilidad en las estimaciones sería recomendable valorar en gran medida la bondad de ajuste que tiene un conjunto de datos a la distribución de probabilidad multivariada.

En ese sentido, requerimos una distribución de probabilidad multivariada que permita capturar los rasgos observados en los rendimientos financieros. A lo largo de los últimos años se ha encontrado que las distribuciones de probabilidad de rendimientos financieros usualmente no se distribuyen de acuerdo a las características de una variable aleatoria

normal, Trejo, Valdés y Núñez (2006). Sin embargo, los inconvenientes numéricos que a menudo presentan las distribuciones no normales multivariadas, limitan su aplicación y extensión hacia los modelos económicos y financieros, Protasov (2004). En esta tesis se incorpora a la familia de distribuciones de probabilidad hiperbólicas generalizadas (GH por sus siglas en inglés) como una herramienta para estimar lo mejor posible la matriz de covarianzas y discernir tanto la variación libre de ruido como las características no normales presentes en un conjunto de datos de rendimientos financieros.

Barndorff-Nielsen (1977) introdujo la familia GH en el estudio de los granos de arena, pero pasaron muchos años antes de las implementaciones numéricas robustas. No fue sino hasta que Blæsild y Sørensen (1992) desarrollaron el programa HYP que la familia GH alcanzó aplicaciones concretas. HYP permitía ajustar distribuciones hiperbólicas mediante máxima verosimilitud, aunque solamente para dimensiones menores o iguales a tres. Prause (1999) analizó exhaustivamente diferentes especificaciones de calibración para la función de log-verosimilitud de GH mediante HYP, sin embargo la estimación con dimensión menor o igual a tres seguía presentando dificultades numéricas.

Eberlein y Keller (1995) ajustaron la distribución hiperbólica univariada para diferentes rendimientos de acciones alemanas bajo una alta precisión. No obstante, el caso multivariado no fue reportado por el elevado tiempo de cómputo y falta de convergencia que se tiene cuando se realiza la estimación directa mediante máxima verosimilitud, por ello durante muchos años aparecieron pocas aplicaciones de la familia GH bajo un enfoque multivariado.

La forma de resolver el problema de estimación llegó mediante el algoritmo EM (expectation-maximization). Este poderoso procedimiento que presentaron Dempster, Laird y Rubin (1977) puede utilizarse en estimaciones de máxima verosimilitud. Liu y Rubin (1995) emplean este algoritmo junto con sus extensiones ECME (expectation

conditional maximization) y MCECM (multicycle expectation conditional maximization) para calibrar la distribución multivariada t-Student.

El algoritmo EM fue modificado por Protasov (2004) para estimar las distribuciones hiperbólicas generalizadas multivariadas con el parámetro λ fijo de la función de Bessel de tercer orden en la función de log-verosimilitud aumentada. Este autor es el primero que reporta la estimación de la familia GH para una dimensión mayor a tres. Específicamente, Protasov (2004) ajusta la distribución normal inversa gaussiana (NIG) de dimensión cinco para distintos rendimientos sobre diferentes tipos de cambio, donde $\lambda = -1/2$.

Hu (2005) analiza exhaustivamente los casos límite de la familia GH y encuentra que el algoritmo de Protasov (2004) es inestable numéricamente para valores pequeños del parámetro $|\lambda| < 1$. Por ello combina apropiadamente la especificación de McNeil, Frey y Embrechts (2005) con el procedimiento de Protasov (2004) para estimar con mayor precisión los casos límite de la familia GH: la distribución t-Student, Varianza Gama y gaussiana inversa normal (NIG, por sus siglas en inglés). Hu (2005) reduce el número de iteraciones y el tiempo de procesamiento para valores de $|\lambda| < 10$, para ello fija convenientemente los parámetros χ , ψ de la función gaussiana inversa generalizada (GIG) asociada, y a diferencia de Protasov (2004) estima otras distribuciones diferentes al caso de la normal inversa gaussiana.

Recientemente, Lüthi y Wolfgang (2011) implementan la estimación de la familia GH bajo diferentes parametrizaciones. La estimación en el caso univariado se realiza directamente mediante máxima verosimilitud, como al principio se llevaba a cabo en Prause (1999) y Blasild y Sørensen (1992). Los cálculos para el caso multivariado emplean el procedimiento MCECM, Meng y Rubin (1993), McNeil, Frey, y Embrechts (2005). En este procedimiento el valor de λ puede no ser fijo y estimarse conjuntamente con el resto de los parámetros. No obstante, las rutinas implementadas por Lüthi y Wolf-

gang (2011) no necesariamente convergen a un punto concreto cuando la distribución GH presenta una dimensionalidad superior a diez. En ese sentido, aunque el método de estimación numérica sea preciso y exacto para un amplio rango de parámetros, dimensiones y tamaños de muestra, existe la posibilidad de reducir aún más el número de iteraciones del algoritmo de estimación, elegir alternativamente el punto inicial y evitar complicaciones de convergencia por la inestabilidad del parámetro libre λ en muestras grandes cuya dimensionalidad sea superior a diez.

En esta tesis se realiza una modificación y extensión al algoritmo EM presentado por Protasov (2004) para estimar los parámetros de la familia de distribuciones GH para un rango más amplio de parámetros, dimensionalidad y tamaños de muestra. Se asume que el parámetro de la función de Bessel modificada de tercer orden no es fijo, y se elige sistemáticamente un punto inicial para los parámetros de GH bajo la distribución de probabilidad gaussiana inversa generalizada (GIG). La elección sistemática de este punto bajo la GIG emplea los resultados que existen sobre invarianza de los parámetros en las combinaciones lineales de la GH, McNeil, Frey, y Embrechts (2005). La bondad de ajuste se realiza mediante una prueba empírica para distribuciones multivariadas continuas, McAssey (2013). Estas modificaciones y extensiones facilitan el ajuste de las distribuciones GH y su potencial aplicación en distintas áreas. En particular la estimación alternativa de la matriz de covarianzas para discernir la variación libre de ruido en un conjunto de rendimientos financieros y refinar la solución óptima de un portafolio bajo el enfoque de Markowitz.

Capítulo 2 Enfoques estadísticos

En las siguientes secciones se presenta la descripción de los dos enfoques estadísticos que abarcamos en esta tesis, tanto el caso no paramétrico con matrices aleatorias, como el caso paramétrico de la distribución hiperbólica generalizada multivariada.

2.1 Matrices aleatorias

La matriz de covarianzas de un conjunto de rendimientos para una serie de activos es una herramienta de gran relevancia, sin embargo en la práctica resulta difícil determinar una estimación confiable, ya que existen contribuciones aleatorias por "ruido", Laloux et al (2000). Es posible proceder mediante métodos paramétricos y no paramétricos para aproximar la matriz de covarianzas poblacional. En esta parte presentamos elementos de la teoría de matrices aleatorias, donde no se requiere estimar una distribución de probabilidad conjunta multivariada.

Sabemos que medir la covarianza entre diferentes variables es relevante para tener una medida de la asociación que existe entre las variables de un cierto portafolio. Si tenemos N activos con precio p_i en el tiempo t , $i = 1, 2, \dots, N$, $t = 0, 1, \dots, T$, entonces el rendimiento geométrico se define como sigue:

$$r_{it} = \ln(p_{it}) - \ln(p_{i,t-1})$$

La matriz de correlaciones empírica puede no reflejar la correcta asociación poblacional entre los diferentes rendimientos, pues el conjunto de activos considerados pueden arrojar información con ruido. En ese sentido, la teoría de matrices aleatorias (TMA) permite clasificar la variación por ruido en una muestra aleatoria mediante distintas pruebas de hipótesis, ya que se justifica que las propiedades de la TMA son aplicables en las matrices de correlación empíricas, Laloux (1999).

El conjunto de matrices aleatorias simétricas C de tamaño $N \times N$, cuyos elementos son variables aleatorias independientes e idénticamente distribuidas se conoce como el conjunto de matrices aleatorias, Mehta (2004), donde $C = (1/T)AA'$, A es una matriz de tamaño $N \times T$ y A' es la matriz transpuesta de A .

Cuando el tamaño de la muestra T tiende a infinito, la distribución de los valores propios de C tiene propiedades universales, que son independientes de la distribución de los elementos de la matriz C , Bouchaud y Potters (2003). Recordemos que λ es un valor propio de la matriz cuadrada C con vector propio asociado $v \neq 0$ si se satisface que

$$Cv = \lambda v.$$

Además existen tantos valores propios $\lambda_1, \lambda_2, \dots, \lambda_T$ y vectores propios v_1, v_2, \dots, v_T como el número de columnas de la matriz cuadrada C . Si v es un vector propio de C también lo será kv para cualquier $k \in \mathbb{R}$. Para evitar esta ambigüedad suponemos que los vectores propios están normalizados de manera que $|v| = 1$. Sin embargo, el signo queda indeterminado: si v es un vector propio también lo es $-v$.

Específicamente, la función de densidad de probabilidad de los valores propios $\lambda_1, \lambda_2, \dots, \lambda_T$, se define como

$$f(\lambda_i) = \frac{Q}{2\pi\sigma^2} \frac{\sqrt{(\lambda_{\text{máx}} - \lambda_i)(\lambda_i - \lambda_{\text{mín}})}}{\lambda_i}$$

donde $Q = T/N \geq 1$ es un cociente fijo y $\lambda_i \in [\lambda_{\min}, \lambda_{\max}]$ para $i \leq n$, véase figura 2.1.

Los valores propios máximo λ_{\max} y mínimo λ_{\min} se definen como

$$\lambda_{\max} = \sigma^2 \left(1 + \frac{1}{Q} + 2\sqrt{\frac{1}{Q}} \right)$$

$$\lambda_{\min} = \sigma^2 \left(1 + \frac{1}{Q} - 2\sqrt{\frac{1}{Q}} \right)$$

donde σ^2 es la varianza de los elementos de la matriz A . De esta forma podemos decidir cuantitativamente cuando un valor propio λ_i de una matriz de correlación empírica representa una señal de información poblacional genuina y no simplemente un efecto de ruido presente en la muestra de datos. En términos prácticos, cualquier valor propio fuera del intervalo $[\lambda_{\min}, \lambda_{\max}]$ puede considerarse como una señal de información real, Laloux et al. (1999).

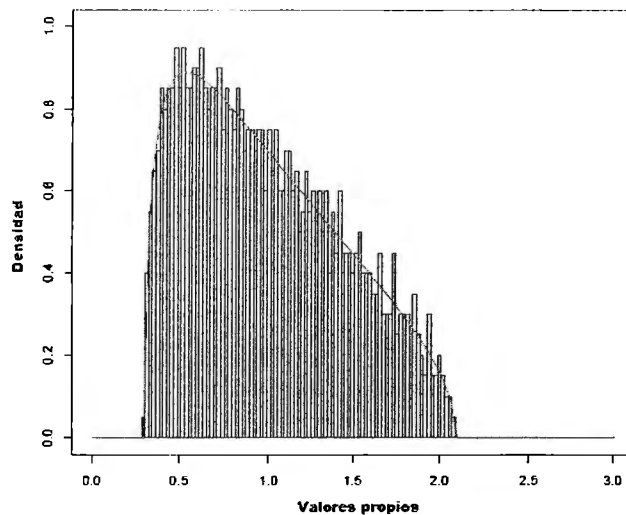


Figura 2.1. Distribución de los valores propios de una matriz aleatoria.

Para verificar si los valores propios dentro del intervalo $[\lambda_{\min}, \lambda_{\max}]$ satisfacen las propiedades universales de las matrices aleatorias, Bai (1999), convertimos los valores propios λ_i en nuevas variables llamadas valores propios desplegados, Giada y Marsili (2001). Los valores propios desplegados tienen un valor promedio uniforme a través del espectro, y se definen como

$$\eta_k = \int_{-\infty}^{\lambda_k} f(x) dx$$

para $k = 1, 2, \dots, T$, Meyer (2008).

En el caso de las matrices aleatorias se cumple que los valores propios desplegados η_k cumplen propiedades específicas:

1.- En particular, la distribución de probabilidad de la diferencia entre vecinos más cercanos, Gupta y Nagar (2000),

$$S_k = \eta_{k+1} - \eta_k$$

está dada por la expresión

$$g(s_k) = \frac{\pi s_k}{2} \exp\left(-\frac{\pi}{4} s_k^2\right).$$

2.- La distribución de probabilidad de la diferencia entre segundos vecinos más cercanos, Giada (2001),

$$U_k = \eta_{k+2} - \eta_k$$

está dada por

$$p(u_k) = \frac{2^{18}}{3^6 \pi^3} u_k^4 \exp\left(-\frac{64}{9\pi} u_k^2\right).$$

3.- Correlaciones entre valores propios desplegados de rango largo.

Para verificar esta propiedad entre pares de valores propios desplegados, se emplea

la “varianza número” $\Gamma = V[f(l)]$, es decir, la varianza del número de valores propios desplegados $f(l)$ en intervalos de longitud l alrededor de cada η_k . En otras palabras, Γ es el número de valores propios desplegados en el intervalo dado por

$$\left[\eta_k - \frac{l}{2}, \eta_k + \frac{l}{2} \right]$$

para $l > 0$. En matrices aleatorias, para valores razonablemente grandes de l , Brody (1981), la varianza número es aproximadamente igual al logaritmo natural de la longitud con valor l .

Si es el caso que se satisfacen las propiedades anteriores es posible encontrar evidencia para rechazar la hipótesis nula de una matriz aleatoria C y por tanto evidencia a favor de que los valores propios fuera del intervalo $[\lambda_{\min}, \lambda_{\max}]$ miden variación libre de ruido, en tanto que los valores propios interiores capturan choques aleatorios. Este resultado es de gran relevancia, pues permite filtrar la matriz de covarianzas al discernir el componente aleatorio atribuible al ruido, Laloux et al (2000).

Si una matriz cuadrada de tamaño N tiene N valores propios distintos, entonces a cada valor propio le podemos asociar un vector propio bien definido y se demuestra que el conjunto de N vectores propios es linealmente independiente, Jhonson (2007). Si un valor propio es múltiple, es decir, la matriz no tiene N valores propios diferentes, los vectores propios asociados a valores propios con multiplicidad mayor a uno no están definidos en general de manera única.

En nuestro contexto, dado que C es una matriz cuadrada simétrica entonces se cumple que tiene valores propios reales y vectores propios ortogonales entre sí. Como consecuen-

cia, existe una matriz ortogonal V y una matriz diagonal D tal que

$$V'CV = D$$

donde V tiene por columnas a los vectores propios y D contiene a los valores propios, Jhonson (2007). Luego la descomposición espectral de la matriz C

$$C = VDV'$$

puede escribirse como

$$C = \sum_{i=1}^N \lambda_i v_i v_i'$$

En otras palabras, la matriz C puede escribirse como suma de N matrices de rango uno $v_i v_i'$ con ponderaciones λ_i . Esta descomposición permite filtrar el ruido según las pruebas de hipótesis de matrices aleatorias, ya que basta remover los valores propios asociados al ruido para eliminar la parte no sistemática de la matriz de covarianzas empírica. Otra parte relevante de esta descomposición es que si algunos valores propios son muy pequeños, podemos evaluar cuál es su impacto cuando realizamos cálculos con la matriz inversa, pues

$$C^{-1} = \sum_{i=1}^N \lambda_i^{-1} v_i v_i'$$

tiene los mismos vectores propios que C y valores propios λ_i^{-1} , $i = 1, 2, \dots, N$.

Por otra parte, para medir el número de componentes que participan significativamente en cada vector propio v_k , se emplea el cociente inverso de participación, Laloux

(1999), el cual se define como el inverso de I_k , donde

$$I_k = \sum_{i=1}^N v_{ki}^4$$

siendo N el número de activos, y por ende el número de componentes.

La j -ésima componente del vector propio v_{ki} corresponde a la contribución del j -ésimo activo al vector propio v_{ki} . Mediante el inverso de I_k es posible medir cuál es el número de activos cuya participación contribuye en mayor medida a la varianza capturada por el valor propio λ_k fuera del intervalo $[\lambda_{min}, \lambda_{max}]$, Burda (2004).

Si es el caso, que todas las componentes son idénticas y $v_{ki} = 1/\sqrt{N}$, entonces

$$\begin{aligned} I_k &= \sum_{i=1}^N v_{ki}^4 \\ &= \sum_{i=1}^N \left(\frac{1}{\sqrt{N}} \right)^4 \\ &= \sum_{i=1}^N \left(\frac{1}{N} \right)^2 \\ &= \frac{1}{N} \end{aligned}$$

y el coeficiente inverso de participación nos indicaría que todos los elementos del vector propio contribuyen de igual forma a la varianza del valor propio. En otro caso extremo, si sólo hubiese un componente diferente de cero, entonces $I_k = 1$ y se concluiría que la variación del valor propio se atribuye totalmente a un sólo componente.

2.2 La Distribución Hiperbólica Generalizada

La estimación de la matriz de covarianzas entre diferentes variables aleatorias resulta más adecuada si consideramos una distribución de probabilidad conjunta multivariada, ya que la matriz de covarianzas estimada puede pensarse como un estimador de la matriz de covarianzas poblacional subyacente. Una distribución de probabilidad multivariada versátil que puede capturar mediante sus parámetros los hechos estilizados de colas gordas, sesgo y picudez de los rendimientos de activos es la familia GH, Protasov (2004).

Una distribución de probabilidad hiperbólica generalizada (GH) consta de distintas partes. Por ello definimos la función de Bessel de tercer orden con parámetro λ , luego mediante ella establecemos la función de densidad inversa gaussiana generalizada (GIG) y definimos finalmente la función de densidad hiperbólica generalizada (GH). Ahora bien, las funciones de distribución de probabilidad hiperbólicas generalizadas presentan diferentes parametrizaciones. Usualmente, la literatura define la función de densidad de probabilidad similarmente a Barndorff-Nielsen (1977), como se presenta directamente en Prause (1999) y Protasov (2004). No obstante, esa definición presenta el inconveniente de que los parámetros no son invariantes bajo transformaciones lineales. En este documento empleamos la especificación de McNeil, Frey, y Embrechts (2005) para trabajar con la familia GH como la mezcla media-varianza entre una variable aleatoria normal multivariada y una variable aleatoria generalizada inversa gaussiana (GIG), y de esa forma utilizar convenientemente la propiedad de invarianza.

2.2.1 La función de Bessel de tercer orden

La función de Bessel de tercer orden puede tener diferentes representaciones, en este trabajo empleamos la especificación integral, Paoletta (2007).

Definición 1. La representación integral de la función modificada de Bessel de tercer orden con parámetro $\lambda \in \mathbb{R}$ para $x > 0$ es

$$K_\lambda(x) = \frac{1}{2} \int_0^\infty t^{\lambda-1} e^{-\frac{1}{2}x(t+t^{-1})} dt$$

Esta función es similar a la función *gamma* y para cualesquiera $\lambda \in \mathbb{R}$, $x \in \mathbb{R}_+$ cumple que

- a) $K_\lambda(x) = K_{-\lambda}(x)$
- b) $K_{\lambda+1}(x) = \frac{2\lambda}{x} K_\lambda(x) + K_{\lambda-1}(x)$

Adicionalmente se define una función auxiliar,

$$k_\lambda(\chi, \psi) = \int_0^\infty x^{\lambda-1} e^{-\frac{1}{2}(\chi x^{-1} + \psi x)} dx.$$

tal que

- a) $k_\lambda(\chi, \psi) = 2\eta^{\lambda/2} K_\lambda(\omega)$
- b) $k_\lambda(0, \psi) = \left(\frac{\psi}{2}\right)^{-\lambda} \Gamma(\lambda)$
- c) $k_\lambda(\chi, 0) = \left(\frac{\chi}{2}\right)^\lambda \Gamma(-\lambda)$
- d) $k_\lambda(\chi, \psi) = k_{-\lambda}(\psi, \chi)$
- e) $k_\lambda(\chi, \psi) = r^\lambda k_\lambda(r^{-1}\chi, r\psi)$

siendo que $\eta = \frac{\chi}{\psi}$, $\omega = \sqrt{\chi\psi}$, $\lambda \in \mathbb{R}$, $\chi, \psi \in \mathbb{R}_+$ y la función gamma definida por

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$$

para $z > 0$. Las funciones K , k y Γ constituyen los bloques fundamentales para definir las funciones de densidad de probabilidad de la familia GH.

2.2.2 Distribución gaussiana inversa generalizada

En esta sección definimos una variable aleatoria gaussiana inversa generalizada (GIG), la cual resulta crucial para establecer las funciones de densidad de la familia GH en forma explícita.

Definición 2. Sea Z una variable aleatoria continua positiva, $\mu, \beta \in \mathbb{R}$. Si X es una variable aleatoria que satisface

$$X|Z \sim N(\mu + \beta Z, Z).$$

Decimos que X es una variable aleatoria *media – varianza*, Paoletta (2007) y la función de densidad de probabilidad es

$$f_X(x) = \int_0^{\infty} f_N(x; \mu + \beta z, z) f_Z(z) dz$$

donde el valor esperado y la varianza son

$$E[X] = \mu + \beta E[Z],$$

$$V[X] = E[Z] + \beta^2 V[Z],$$

respectivamente.

La función generadora de momentos y la función característica son

$$m_X(t) = e^{\mu t} m_Z(\beta t + t^2/2),$$

$$\varphi_X(v) = e^{i\mu v} \varphi_Z(\beta v + iv^2/2).$$

Definición 3. Decimos que una variable aleatoria X es inversa gaussiana generalizada (GIG) con parámetros $\lambda \in \mathbb{R}$, $\chi \geq 0$ y $\psi \geq 0$ si la función de densidad de probabilidad, Paoletta (2007), es

$$f_X(x; \lambda, \chi, \psi) = \frac{1}{k_\lambda(\chi, \psi)} x^{\lambda-1} \exp \left[-\frac{1}{2} (\chi x^{-1} + \psi x) \right]$$

para $x \in \mathbb{R}_+$ y escribimos $X \sim GIG(\lambda, \chi, \psi)$. A saber existen tres casos posibles según el dominio de variación de sus parámetros.

- a) Caso normal: $\lambda \in \mathbb{R}$, $\chi > 0$ y $\psi > 0$.
- b) Caso frontera I: $\lambda > 0$, $\chi = 0$ y $\psi > 0$.
- c) Caso frontera II: $\lambda > 0$, $\chi > 0$ y $\psi = 0$.

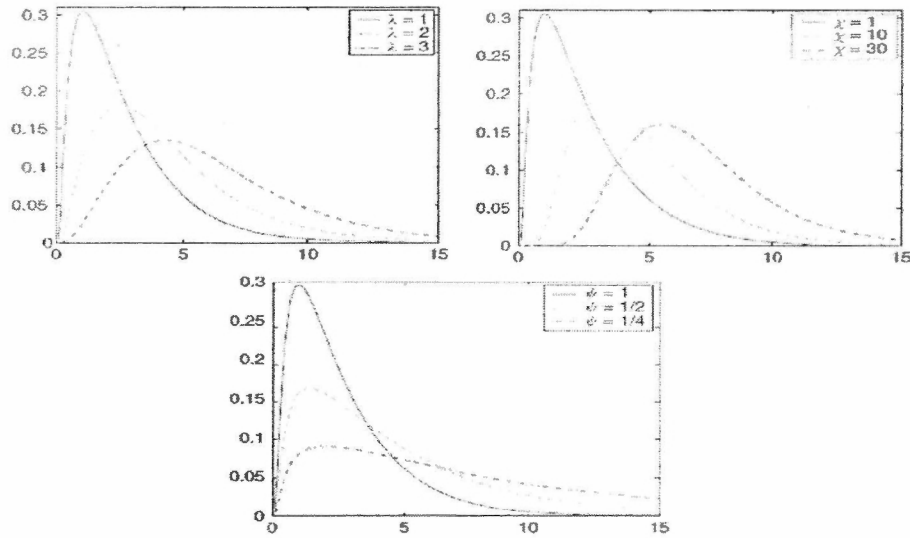


Figura 2.2. GIG con parámetros no especificados igual a la unidad.

Además se tiene que el valor esperado es

$$E[X] = \frac{k_{\lambda+1}(\chi, \psi)}{k_{\lambda}(\chi, \psi)}$$

y la varianza es

$$V[X] = \frac{k_{\lambda}(\chi, \psi)k_{\lambda+2}(\chi, \psi) - [k_{\lambda+1}(\chi, \psi)]^2}{[k_{\lambda}(\chi, \psi)]^2}.$$

El r -ésimo momento de la variable aleatoria X es

$$E[X^r] = \int_0^{\infty} x^r f_X(x; \lambda, \chi, \psi) dx = \eta^r \frac{K_{\lambda+r}(\omega)}{K_{\lambda}(\omega)}.$$

Más aún se tiene que

$$E[\ln(X)] = \ln(\eta) + \frac{\partial}{\partial \lambda} \frac{K_{\lambda}(\omega)}{K_{\lambda}(\omega)}.$$

Esta última expresión resulta muy útil para realizar las estimaciones numéricas que se presentan más adelante.

La función generadora de momentos es

$$m_{GIG}(t; \lambda, \chi, \psi) = \frac{k_{\lambda}(\chi, \psi - 2t)}{k_{\lambda}(\chi, \psi)}.$$

Adicionalmente, debemos comentar que dependiendo de los casos límite planteados arriba tenemos la siguiente subfamilia para la variable aleatoria $GIG(\lambda, \chi, \psi)$, cuadro 2.1. El comportamiento de la variable GIG según la variación de sus parámetros determinara en

gran medida y forma la función de densidad de probabilidad de la familia GH.

Cuadro 2.1. Casos especiales de la GIG.

Nombre	Rango de parámetros		
normal	$\lambda \in \mathbb{R}$	$\chi > 0$	$\psi > 0$
gamma	$\lambda > 0$	$\chi = 0$	$\psi > 0$
inversa gamma	$\lambda < 0$	$\chi > 0$	$\psi = 0$
exponencial	$\lambda = 1$	$\chi = 0$	$\psi > 0$
hipérbolica positiva	$\lambda = 1$	$\chi > 0$	$\psi > 0$
Lévy	$\lambda = -1/2$	$\chi > 0$	$\psi = 0$
inversa gaussiana	$\lambda = -1/2$	$\chi > 0$	$\psi > 0$
caso límite de Dirac, $a > 0$	$\lambda \in \mathbb{R}$	$\eta \rightarrow a$	$\omega \rightarrow \infty$

2.2.3 Distribución hiperbólica generalizada univariada

La distribución hiperbólica generalizada univariada se define como una mezcla media-varianza, Paoella (2007), donde el peso es una variable aleatoria $GIG(\lambda, \chi, \psi)$.

Definición 4. Sean $\alpha, \delta \geq 0$ y $\lambda, \mu, \beta \in \mathbb{R}$ tales que $\beta \in [-\alpha, \alpha]$, decimos que X sigue una distribución hiperbólica generalizada univariada si la función de densidad de probabilidad asociada es

$$f_X(x; \lambda, \alpha, \beta, \delta, \mu) = \frac{k_{\lambda-1/2}((x-\mu)^2 + \delta^2 \cdot \alpha^2)}{\sqrt{2\pi} k_\lambda(\delta^2, \alpha^2 - \beta^2)} \exp[\beta(x-\mu)],$$

donde la variable aleatoria implícita en media-varianza es una variable aleatoria $GIG(\lambda, \chi, \psi)$

con

$$\begin{aligned} \chi &= \delta^2 \\ \psi &= \alpha^2 - \beta^2. \end{aligned}$$

El valor esperado es

$$E[X] = \mu + \beta \frac{k_{\lambda+1}(\chi, \psi)}{k_{\lambda}(\chi, \psi)}$$

y la varianza es

$$V[X] = \frac{k_{\lambda+1}(\chi, \psi)}{k_{\lambda}(\chi, \psi)} + \beta^2 \frac{k_{\lambda}(\chi, \psi)k_{\lambda+2}(\chi, \psi) - [k_{\lambda+1}(\chi, \psi)]^2}{[k_{\lambda}(\chi, \psi)]^2}.$$

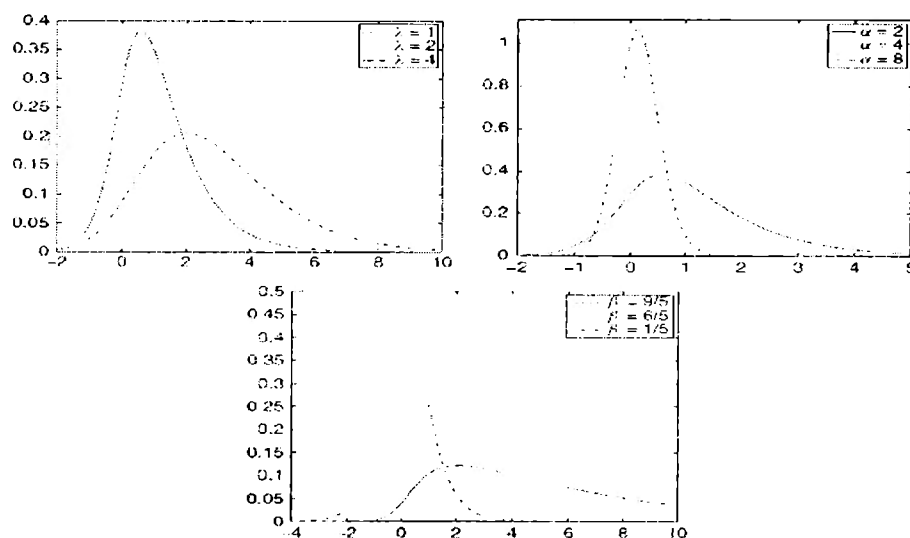


Figura 2.3. La GH univariada para $\delta = 1$ y $\mu = 0$: a) $\alpha = 2$ y $\beta = 1$, b) $\lambda = 1$ y $\beta = 1$,
 c) $\lambda = 1$ y $\alpha = 2$.

La función generadora de momentos es

$$m_X(t) = e^{\mu t} \frac{k_{\lambda}(\delta^2, \alpha^2 - (\beta + t)^2)}{k_{\lambda}(\delta^2, \alpha^2 - \beta^2)}$$

donde $-\alpha - \beta < t < \alpha - \beta$.

Intuitivamente podemos interpretar los diferentes parámetros como sigue:

- a) α es un parámetro de cola, pues su magnitud determina el peso de las colas en la distribución.
- b) β es el parámetro de sesgo, si su valor es cero, la distribución es simétrica.
- c) μ es un parámetro de localización y si es el caso que $\beta = 0$, entonces μ coincide con la media.
- d) δ es un parámetro de picudez y controla la forma de la distribución alrededor de la moda.
- e) λ es un parámetro de forma que influye de manera general en la dispersión entre ambas colas de la distribución.

En el cuadro 2.2 presentamos a la subfamilia GHYP univariada.

Cuadro 2.2. Casos especiales de la GHYP.

Nombre	Rango de parámetros				
varianza-gamma	$\lambda > 0$	$\alpha > 0$	$\beta \in (-\alpha, \alpha)$	$\delta = 0$	$\mu \in \mathbb{R}$
Laplace asimétrica	$\lambda = 1$	$\alpha > 0$	$\beta \in (-\alpha, \alpha)$	$\delta = 0$	$\mu \in \mathbb{R}$
Laplace	$\lambda = 1$	$\alpha > 0$	$\beta = 0$	$\delta = 0$	$\mu \in \mathbb{R}$
hiperbólica	$\lambda = 1$	$\alpha > 0$	$\beta \in (-\alpha, \alpha)$	$\delta > 0$	$\mu \in \mathbb{R}$
t asimétrica hiperbólica	$\lambda > 0$	$\alpha = \beta $	$\beta \geq 0$	$\delta > 0$	$\mu \in \mathbb{R}$
t-student	$\lambda > 0$	$\alpha = 0$	$\beta = 0$	$\delta > 0$	$\mu \in \mathbb{R}$
Cauchy	$\lambda = -1/2$	$\alpha = 0$	$\beta = 0$	$\delta > 0$	$\mu \in \mathbb{R}$
gaussiana inversa normal	$\lambda = -1/2$	$\alpha > 0$	$\beta \in (-\alpha, \alpha)$	$\delta > 0$	$\mu \in \mathbb{R}$
normal	$\lambda \in \mathbb{R}$	$\alpha \rightarrow \infty$	$\beta \rightarrow \beta_0$	$\frac{\delta}{\alpha} \rightarrow \sigma^2$	$\mu \in \mathbb{R}$

2.2.4 Función de distribución hiperbólica generalizada multivariada

La distribución hiperbólica generalizada multivariada puede representarse como una mezcla media-varianza entre una variable aleatoria normal multivariada y una variable aleatoria GIG, McNeil, Frey y Embrechts (2005). Esta especificación se establece en términos de los parámetros λ, χ, ψ y los parámetros de una distribución normal multivariante con media μ , varianza Σ y un vector de sesgo γ .

Definición 5. Un vector aleatorio $X \in \mathbb{R}^m$ se dice que tiene una distribución media-varianza normal multivariada si

$$X = \mu + W\gamma + \sqrt{W}AZ,$$

donde

- a) $Z \sim N_m(0, I_m)$
- b) W es una variable aleatoria escalar, positiva e independiente de Z ,
- c) A es una matriz de constantes de tamaño $m \times m$,
- d) μ, γ son vectores en \mathbb{R}^m .

Notemos que en ese caso se tiene que

$$X|W \sim N_m(\mu + W\gamma, W\Sigma)$$

siendo

$$\Sigma = AA'.$$

La variable aleatoria W tiene varianza finita y puede pensarse como un shock que con-

tinuamente cambia la media y la varianza de un vector aleatorio. Si la variable aleatoria W se distribuye como una variable aleatoria $GIG(\lambda, \chi, \psi)$, entonces decimos que X se distribuye como una hiperbólica generalizada multivariada. Los parámetros λ, χ, ψ determinan que tanto peso se asigna a las colas y al centro de la distribución. En general, mientras más grandes son estos parámetros, mayor cercanía existe a la distribución normal. En tanto que μ es un vector de localización, $\Sigma = AA'$ es la matriz de dispersión y γ es un vector de sesgo. El parámetro γ mide el sesgo respecto de la variable aleatoria normal de la mezcla, así que si su valor es cero, entonces la variable X es simétrica, Paoletta (2007).

Además se cumple que

$$\begin{aligned} E[X] &= \mu + E[W]\gamma, \\ V[X] &= E[W]\Sigma + V[W]\gamma\gamma'. \end{aligned}$$

En particular, para el caso de la función de densidad de probabilidad hiperbólica generalizada tenemos el siguiente resultado.

Teorema 1 (McNeil, Frey y Embrechts (2005)). Si la variable de mezclado W es $GIG(\lambda, \chi, \psi)$, entonces la función de densidad de probabilidad de X es

$$f_X(x; \lambda, \chi, \psi, \mu, \Sigma, \gamma) = c \frac{K_{\lambda - \frac{m}{2}} \left(\sqrt{(\chi + (x - \mu)' \Sigma^{-1} (x - \mu)) (\psi + \gamma' \Sigma^{-1} \gamma)} \right) e^{(x - \mu)' \Sigma^{-1} \gamma}}{\left(\sqrt{(\chi + (x - \mu)' \Sigma^{-1} (x - \mu)) (\psi + \gamma' \Sigma^{-1} \gamma)} \right)^{\frac{m}{2} - \lambda}},$$

donde Σ tiene rango m y la constante de normalización es

$$c = \frac{(\sqrt{\chi\psi})^{-\lambda} \psi^\lambda (\psi + \gamma' \Sigma^{-1} \gamma)^{\frac{m}{2} - \lambda}}{(2\pi)^{\frac{m}{2}} |\Sigma|^{\frac{1}{2}} K_\lambda(\sqrt{\chi\psi})}.$$

Existe un problema de identificación en los parámetros que determinan la función de densidad de probabilidad, Hu (2005), ya que para cualquier $\theta > 0$

$$f_X(x; \lambda, \chi/\theta, \theta\psi, \mu, \theta\Sigma, \theta\gamma) = f_X(x; \lambda, \chi, \psi, \mu, \Sigma, \gamma).$$

Este problema causa grandes inconvenientes durante la estimación y calibración de la distribución hiperbólica generalizada, ya que pueden existir problemas de inestabilidad y convergencia, Hu (2005). Protasov (2004) resuelve este problema normalizando el determinante de Σ a la unidad, a esta nueva matriz de dispersión la denotamos por Δ ,

$$\Delta = \frac{\Sigma}{|\Sigma|}.$$

McNeil, Frey y Embrechts (2005) emplean la especificación inicial de Barndorff-Nielsen (1977)

$$f_X(x; \lambda, \alpha, \mu, \Delta, \delta, \beta) = \frac{(\alpha^2 - \beta' \Delta \beta)^{\lambda/2}}{(2\pi)^{\frac{n}{2}} \sqrt{|\Delta|} \alpha^{\lambda - \frac{m}{2}} \delta^\lambda K_\lambda \left(\delta \sqrt{\alpha^2 - \beta' \Delta \beta} \right)} \times \frac{K_{\lambda - \frac{m}{2}} \left(\alpha \sqrt{\delta^2 + (x - \mu)' \Delta^{-1} (x - \mu)} \right) e^{\beta'(x - \mu)}}{\left(\sqrt{\delta^2 + (x - \mu)' \Delta^{-1} (x - \mu)} \right)^{\frac{m}{2} - \lambda}}$$

con la normalización citada por Protasov (2004) para mejorar la precisión y exactitud de un algoritmo de estimación.

Más aún notemos que

$$\begin{aligned} X|W &\sim N_m(\mu + w\beta\Delta, w\Delta) \\ W &\sim GIG(\lambda, \delta^2, \alpha^2 - \beta'\Delta\beta) \\ \chi &= \delta^2 \\ \psi &= \alpha^2 - \beta'\Delta\beta \end{aligned}$$

y que la función generadora de momentos es

$$m_X(t) = e^{t'\mu} \left(\frac{\psi}{\psi - 2t'\gamma - t'\Sigma t} \right)^{\lambda/2} \frac{K_\lambda \left(\sqrt{\psi (\chi - 2t'\gamma - t'\Sigma t)} \right)}{K_\lambda(\sqrt{\chi\psi})}$$

con $\chi \geq 2t'\gamma + t'\Sigma t$, Hu (2005). En este contexto tendríamos para el caso univariado que la función de densidad de probabilidad es

$$f(x) = \frac{(\alpha^2 - \beta^2)^{\lambda/2}}{\sqrt{2\pi}\alpha^{\lambda-1/2}\delta^k K_\lambda(\delta\sqrt{\alpha^2 - \beta^2})} \frac{K_{\lambda-1/2}}{\left(\sqrt{\delta^2 + (x - \mu)^2}\right)^{1/2-\lambda}} \exp(\beta(x - \mu))$$

Esta última expresión resultara relevante, ya que mediante el teorema de McNeil, Frey y Embrechts (2005) puede caracterizarse una combinación de los elementos del vector aleatorio X de la familia GH. En ese sentido, es importante comentar que existen diferentes parametrizaciones, las cuales pueden ser útiles numéricamente según los parámetros que sean relevantes, Lüthi (2011).

Capítulo 3

Estimación de la distribución hiperbólica generalizada multivariada

Los modelos que emplean la distribución hiperbólica generalizada univariada han sido ampliamente comentados en la literatura, Prause (1999). Sin embargo, el caso multivariado ha encontrado menos aplicaciones, pues maximizar la función de verosimilitud directamente presenta dificultades numéricas, Hu (2005). Protasov (2004) resuelve parcialmente este problema mediante el algoritmo EM, pues fija el parámetro de la función de Bessel en $\lambda = -1/2$ y estima la NIG multivariada de dimensión cinco. Asimismo, Hu (2005) fija convenientemente los parámetros χ, ψ y estima las distribuciones varianza Gamma y t-Student generalizada para diferentes valores fijos de λ . En esta sección revisamos el algoritmo de Protasov (2004) bajo la especificación de McNeil, Frey y Embrechts (2005) para implementar el algoritmo EM y estimar convenientemente a los miembros de la familia GH con el parámetro λ libre.

3.1 El algoritmo de estimación EM

La especificación univariada de la familia GH ha sido ampliamente utilizada, no así la distribución hiperbólica generalizada multivariada debido a algunas complicaciones numéricas. La representación media-varianza permite emplear fácilmente el algoritmo EM, Dempster et al (1977). En esta sección describimos el algoritmo descrito en Protasov (2004) y McNeil, Frey y Embrechts (2005) junto con las modificaciones que consideramos pertinentes para ampliar su uso en el caso del parámetro λ libre de la función de Bessel de tercer orden.

Consideremos una muestra de vectores aleatorios $X_1, X_2, \dots, X_m \in \mathbb{R}^n$, independientes e idénticamente distribuidos provenientes de una población que se distribuye bajo una distribución hiperbólica generalizada multivariada cuyos parámetros son $\Theta = (\lambda, \gamma, \delta, \beta, \mu, \Delta)$. Entonces la función de log-verosimilitud es

$$\ln(L(\Theta; x_1, x_2, \dots, x_m)) = \sum_{i=1}^m \ln[f_X(\Theta; x_i)]$$

La estimación mediante máxima verosimilitud directa de los parámetros de interés presenta dificultades numéricas de convergencia cuando la dimensión es mayor a tres, Hu (2005). El punto clave del algoritmo EM para evitar en la medida de lo posible este problema es que introduce variables latentes w_1, w_2, \dots, w_m cuya distribución GIG implica una función de log-verosimilitud aumentada, Lüthi (2011),

$$\ln(L'(\Theta; x_1, x_2, \dots, x_m, w_1, w_2, \dots, w_m)) = \sum_{i=1}^m \ln(f_{X,W}(x_i, w_i; \Theta))$$

Debido a la definición media-varianza de la distribución GH, la función de log-verosimilitud puede reescribirse como

$$\begin{aligned}
\ln(L'(\Theta; x_1, x_2, \dots, x_m, w_1, w_2, \dots, w_m)) &= \sum_{i=1}^m \ln(f_{X|W}(x_i|w_i; \mu, \Delta, \beta)) + \\
&\quad \sum_{i=1}^m \ln(f_W(w_i; \lambda, \gamma, \delta)) \\
&= L_1(\mu, \Delta, \beta; x_1, x_2, \dots, x_m|w_1, w_2, \dots, w_m) + \\
&\quad L_2(\lambda, \gamma, \delta; w_1, w_2, \dots, w_m),
\end{aligned}$$

donde

$$X|W \sim N(\mu + w\beta\Delta, w\Delta),$$

$$W \sim GIG(\lambda, \gamma, \delta)$$

$f_W(w_i; \lambda, \gamma, \delta)$ es la función de densidad de probabilidad normal multivariante condicional y $f_W(w_i; \lambda, \gamma, \delta)$ es la función de densidad de probabilidad GIG de la variable asociada en la mezcla. Notemos que la estimación puede realizarse separadamente para las funciones L_1 y L_2 , respectivamente.

Si la función de densidad de probabilidad, Hu (2005), normal multivariante condicional es

$$f_{X|W}(x|w) = \frac{1}{(2\pi)^{m/2} |\Delta|^{1/2} w^{m/2}} \exp((x - \mu)' \Delta^{-1} \beta \Delta) \exp\left(\frac{-\rho}{2w}\right) \exp\left(-\frac{w}{2} \Delta' \beta' \Delta^{-1} \beta \Delta\right)$$

donde

$$\rho = (x - \mu)' \Delta^{-1} (x - \mu).$$

Entonces la función de log-verosimilitud L_1 es

$$L_1 = -\frac{m}{2} \ln(|\Delta|) - \frac{m}{2} \sum_{i=1}^m \ln(w_i) + \sum_{i=1}^m (x_i - \mu)' \Delta^{-1} \beta \Delta \\ - \frac{1}{2} \sum_{i=1}^m \frac{\rho_i}{w_i} - \frac{1}{2} \Delta' \beta' \Delta^{-1} \beta \Delta \sum_{i=1}^m w_i$$

y la función de log-verosimilitud L_2 es

$$L_2 = (\lambda - 1) \sum_{i=1}^m \ln(w_i) - \frac{\delta^2}{2} \sum_{i=1}^m w_i^{-1} - \frac{\gamma^2}{2} \sum_{i=1}^m w_i - n \lambda \ln(\delta) \\ + n \lambda \ln(\gamma) - n \ln(2K_\lambda(\delta\gamma))$$

En cada caso, si derivamos e igualamos a cero obtenemos de L_1

$$\hat{\beta} = \frac{1}{n} \frac{\sum_{i=1}^m w_i^{-1} (\bar{x} - x_i)}{(\sum_{i=1}^m w_i) (\sum_{i=1}^m w_i^{-1}) - 1} \\ \hat{\mu} = \frac{\sum_{i=1}^m w_i^{-1} x_i - n \hat{\beta}}{\sum_{i=1}^m w_i^{-1}} \\ \hat{\Delta} = \frac{1}{n} \sum_{i=1}^m w_i^{-1} (x_i - \hat{\mu})(x_i - \hat{\mu})' - \frac{1}{n} \sum_{i=1}^m w_i \hat{\beta} \hat{\beta}'$$

En la maximización de L_2 usualmente se asume el parámetro λ como una constante, Hu (2005), y dada la representación integral de la función de Bessel de tercer orden, se resuelve numéricamente para γ y δ . En nuestro caso no conocemos las variables latentes w_i y el parámetro λ no es fijo.

Debido a esta situación, el algoritmo recursivo EM, Lüthi (2011), inicia con una aproximación inicial Θ_0 , el cual consiste básicamente de dos etapas:

Paso E. (Expectation) Calculamos la esperanza condicional de la función de verosimilitud dada la muestra x_1, x_2, \dots, x_n y una estimación puntual $\Theta^{[k]}$ para definir la función

objetivo

$$g(\Theta; \Theta^{[k]}) = E[\ln(L'(\Theta; x_1, x_2, \dots, x_m, w_1, w_2, \dots, w_m)) | x_1, x_2, \dots, x_m; \Theta^{[k]}]$$

Esta parte se apoya sustancialmente de las propiedades de la función de densidad de probabilidad GIG, ya que se requiere estimar $E[W_i | x_i; \Theta^{[k]}]$, $E[W_i^{-1} | x_i; \Theta^{[k]}]$ y $E[\ln(W_i) | x_i; \Theta^{[k]}]$.

Sabemos que

$$f_{W|x}(w|x; \Theta) = \frac{f(x|w; \Theta)h(w; \Theta)}{f(x; \Theta)},$$

entonces, Protasov (2004),

$$W_i | X_i \sim GIG\left(\lambda - \frac{m}{2}, \rho_i + \delta^2, \gamma^2 + \Delta' \beta' \Delta^{-1} \beta \Delta\right).$$

Paso M. (Maximization) Maximizamos la función objetivo $g(\Theta; \Theta^{[k]})$ con respecto a Θ para obtener el siguiente conjunto de parámetros estimados $\Theta^{[k+1]}$.

Existen diferentes variantes, Lüthi (2011), para alternar recursivamente los pasos E y M, McNeil, Frey y Embrechts (2005). El algoritmo MCECM (multi-cycle, expectation, conditional estimation), McNeil, Frey y Embrechts (2005), toma como valor inicial para μ la media muestral, para Δ la matriz de covarianzas empírica y para β un vector de ceros. Se asume el parámetro λ como fijo y se reparametriza la función de densidad de probabilidad de la GH bajo

$$f_X(x; \lambda, \alpha, \mu, \Delta, \delta, \beta) = \frac{(\alpha^2 - \beta' \Delta \beta)^{\lambda/2}}{(2\pi)^{\frac{n}{2}} \sqrt{|\Delta|} \alpha^{\lambda - \frac{n}{2}} \delta^\lambda K_\lambda(\delta \sqrt{\alpha^2 - \beta' \Delta \beta})} \times \frac{K_{\lambda - \frac{n}{2}}\left(\alpha \sqrt{\delta^2 + (x - \mu)' \Delta^{-1} (x - \mu)}\right) e^{\beta'(x - \mu)}}{\left(\sqrt{\delta^2 + (x - \mu)' \Delta^{-1} (x - \mu)}\right)^{\frac{n}{2} - \lambda}}$$

con $\bar{\alpha} = \delta \beta \Delta$, $\gamma = \alpha^2 - \beta' \Delta \beta$. Se puede probar, Lüthi (2011), que

$$\begin{aligned} \delta^2 &= \bar{\alpha} \frac{K_{\lambda+1}(\bar{\alpha})}{K_\lambda(\bar{\alpha})} \\ \gamma^2 &= \bar{\alpha} \frac{K_\lambda(\bar{\alpha})}{K_{\lambda+1}(\bar{\alpha})} \end{aligned}$$

luego el procedimiento MCECM realiza los siguientes pasos:

1. Selecciona un valor inicial $\Theta^{[k]}$.
2. Se calcula $\delta^{[k]}$ y $\gamma^{[k]}$ como funciones de $\bar{\alpha}^{[k]}$.
3. Se estiman los pesos

$$\begin{aligned} \eta^{[k]} &\cong E[W_i | x_i; \Theta^{[k]}], \\ \tau^{[k]} &\cong E[W_i^{-1} | x_i; \Theta^{[k]}] \end{aligned}$$

debido, Protasov (2004), a que

$$W_i | X_i \sim GIG\left(\lambda - \frac{m}{2}, \delta, \gamma\right).$$

4. Se actualizan las estimaciones

$$\begin{aligned}\gamma^{[k+1]} &= \frac{1}{n} \frac{\sum_{i=1}^m \delta_i^{[k]} (\bar{x} - x_i)}{\eta^{[k]} \tau^{[k]} - 1} \\ \mu^{[k+1]} &= \frac{1}{n} \frac{\sum_{i=1}^m \delta_i^{[k]} x_i - \gamma^{[k+1]}}{\tau^{[k]}} \\ \Delta^{[k+1]} &= \frac{1}{n} \sum_{i=1}^m \delta_i^{[k]} (x_i - \mu^{[k+1]}) (x_i - \mu^{[k+1]})' - \eta^{[k]} \gamma^{[k+1]} \gamma^{[k+1]}'\end{aligned}$$

5. Mediante $\bar{\alpha}^{[k]}$ y los valores del punto 4 se actualizan los valores $\eta^{[k]}$ y $\tau^{[k]}$ para luego maximizar la función de log-verosimilitud L_2 y retornar al punto 2 hasta converger. En este último puede dejarse libre el parámetro de la función de Bessel y estimar todos los parámetros, Hu (2005), pero el algoritmo se vuelve inestable para valores de $|\lambda| > 1$.

En este trabajo seguimos el cambio de variable señalado por Protasov (2004), conjuntamente con la definición de McNeil, Frey y Embrechts (2005) para evitar problemas de especificación:

$$\begin{aligned}\tilde{\mu} &= \mu \\ \tilde{\gamma} &= \gamma \\ \tilde{\Sigma} &= \delta^2 \Delta \\ \tilde{\beta} &= \delta^2 \beta \Delta \\ |\Delta| &= 1\end{aligned}$$

Luego manteniendo el parámetro λ fijo tenemos

$$w_i | x_i, \Theta^{[k+1]} \sim GIG(\lambda^*, \delta_i^*, \gamma_{k-1}^*)$$

para $i = 1, 2, \dots, m$ y

$$\begin{aligned}\lambda^* &= \lambda - n/2 \\ \delta_i^* &= \sqrt{1 + (x_i - \tilde{\mu}_{k-1}) \left[\tilde{\Sigma}_{k-1}^{-1} \right] (x_i - \tilde{\mu}_{k-1})'} \\ \gamma_{k-1}^* &= \sqrt{(\tilde{\gamma}_{k-1})^2 + \left(\tilde{\beta}_{k-1} \right) \left[\tilde{\Sigma}_{k-1}^{-1} \right] \left(\tilde{\beta}_{k-1} \right)'}\end{aligned}$$

Los momentos condicionales correspondientes se calculan de acuerdo a $GIG(\lambda^*, \delta_i^*, \gamma_k^*)$:

$$\begin{aligned}E [w_i | x_i, \Theta^{[k+1]}] &= \frac{\delta_i^* K_{\lambda^*+1}(\delta_i^* \gamma_k^*)}{\gamma_k^* K_{\lambda^*}(\delta_i^* \gamma_k^*)} \\ E [w_i^{-1} | x_i, \Theta^{[k+1]}] &= \frac{\gamma_k^* K_{\lambda^*-1}(\delta_i^* \gamma_k^*)}{\delta_i^* K_{\lambda^*}(\delta_i^* \gamma_k^*)}\end{aligned}$$

Las estimaciones de los valores esperados anteriores permiten estimar

$$\tilde{\Sigma}_k = \left(\frac{1}{m} \right) \sum_{i=1}^m \left(\begin{array}{l} (x_i - \tilde{\mu}_k)' (x_i - \tilde{\mu}_k) E [w_i^{-1} | x_i, \Theta^{[k+1]}] - \left(\tilde{\beta}_k \right)' (x_i - \tilde{\mu}_k) \\ - (x_i - \tilde{\mu}_k)' \left(\tilde{\beta}_k \right) + \left(\tilde{\beta}_k \right)' \left(\tilde{\beta}_k \right) E [w_i | x_i, \Theta^{[k+1]}] \end{array} \right)$$

Para actualizar el valor de $\tilde{\gamma}_{k-1}$ se maximiza la función, Protasov (2004),

$$m\lambda \ln(\tilde{\gamma}_{k-1}) - m \ln(K_{\lambda-1}(\tilde{\gamma}_{k-1})) - \left(\frac{\tilde{\gamma}_{k-1}}{2} \right)^2 \sum_{i=1}^m E [w_i | x_i, \Theta^{[k+1]}]$$

y finalmente obtenemos,

$$\begin{aligned}\mu &= \tilde{\mu} \\ \delta &= \left| \tilde{\Sigma} \right|^{1/2n} \\ \gamma &= \tilde{\gamma} / \delta \\ \beta &= \tilde{\beta} \tilde{\Sigma}^{-1} \\ \Delta &= \frac{1}{\left| \tilde{\Sigma} \right|^{1/n}} \tilde{\Sigma}\end{aligned}$$

3.2 El algoritmo EM modificado

La función de densidad univariada adquiere importancia en la estimación de la familia GH multivariada mediante el siguiente resultado.

Proposición 1 (McNeil, Frey y Embrechts (2005)). Si $X \sim GH(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ m -dimensional y $Y = BX + b$, donde B es una matriz de tamaño $k \times m$ y $b \in \mathbb{R}^k$, entonces $Y \sim GH(\lambda, \chi, \psi, B\mu + b, B\Sigma B', B\gamma)$ k -dimensional.

En particular para una combinación lineal de los componentes de X se tiene que la suma ponderada de variables aleatorias hiperbólicas generalizadas $Y = w'X$ sigue una distribución hiperbólica generalizada univariada, pues basta colocar $b = 0$ y $B = w'$ en la proposición anterior para $w = (w_1, w_2, \dots, w_m)$, y en ese caso

$$Y \sim GH(\lambda, \chi, \psi, w'\mu, w\Sigma w', w\gamma).$$

La proposición anterior nos permite obtener un punto inicial sistemático para la ejecución del algoritmo EM en el cálculo de los parámetros de la distribución multivariada GH. Dada la invarianza de los parámetros (λ, χ, ψ) , podemos observar que dada

cualquier combinación lineal de los elementos del vector aleatorio X , debe cumplirse que los parámetros estimados de la GIG implícita son los mismos, si bien es cierto que una muestra aleatoria puede contener diferentes componentes ruidosos, es posible tener un estimador robusto para los parámetros poblacionales para una GH multivariada. El estimador que empleamos es la mediana de todos los valores óptimos $(\hat{\lambda}, \hat{\chi}, \hat{\psi})$ que resultan bajo máxima verosimilitud directa para una distribución hiperbólica generalizada univariada. Las combinaciones lineales que tomamos en cuenta son aquellos vectores B_j que constituyen la base canónica de \mathbb{R}^m . La estimación $(\hat{\lambda}_0, \hat{\chi}_0, \hat{\psi}_0)$ la tomamos como parte del punto inicial para el algoritmo EM presentado en Protasov (2004), así que concretamente los parámetros de la GIG que tomamos como punto de partida es $(\hat{\lambda}_0, \hat{\gamma}_0, \hat{\delta}_0)$, donde $\hat{\delta}_0 = \sqrt{\hat{\chi}_0}$ y $\hat{\gamma}_0 = \sqrt{\hat{\psi}_0}$.

Dado que consideramos el parámetro λ de la función modificada de Bessel de tercer orden sin restricciones, el procedimiento numérico que hemos empleado para ubicar el máximo de la función de verosimilitud no recurre a derivadas. En lugar de ello, se implementa adecuadamente el algoritmo simplex general de Kelley (1999). Este algoritmo se basa en el procedimiento de minimización de una función de n variables, Nelder (1965).

Kelley (1999) considera $n + 1$ puntos iniciales p_1, p_2, \dots, p_n arbitrarios que definen un polígono alrededor de un punto inicial $p_0 \in \mathbb{R}^n$, este punto propuesto es el inicio de las iteraciones hacia el valor mínimo de la función. Se define y_i como el valor de la función en el punto p_i y

$$y_h = \text{máx} \{y_i\}$$

$$y_l = \text{mín} \{y_i\}$$

como los valores alto y bajo, respectivamente. Luego definimos el centroide \bar{p} del polígono

para todos los puntos $i \neq h$ y la distancia entre los puntos p_i, p_j como $[p_i, p_j]$. En cada iteración, p_h se reemplaza por un nuevo punto mediante tres procesos: reflexión, contracción y expansión.

La reflexión de p_h se denota por

$$p^* = (1 + \alpha)\bar{p} - \alpha p_h$$

donde α es una constante positiva llamada coeficiente de reflexión. De esa forma p^* se encuentra en el segmento que une a \bar{p} y p_h . Si y^* se encuentra entre y_l y y_h , entonces p_h se reemplaza por p^* repetimos el procedimiento.

Si $y^* < y_l$, entonces definimos

$$p^{**} = \gamma p^* + (1 - \gamma)\bar{p}$$

donde $\gamma > 1$ se denomina factor de expansión. Si $y^{**} < y_l$ reemplazamos p_h por p^{**} y reiniciamos el proceso, pero si $y^{**} > y_l$ reemplazamos p_h por p^* antes de reiniciar el procedimiento. Además, si cuando realizamos la reflexión de p a p^* encontramos que $y^* > y_l$ para cada $i \neq h$, entonces definimos un nuevo p_h según el valor más pequeño entre y y y^* , y definimos

$$p^{**} = \beta p_h + (1 - \beta)\bar{p}$$

siendo $0 < \beta < 1$ un coeficiente de contracción. Tomando p^{**} continuamos el proceso, salvo que $y^{**} > \min\{y_h, y^*\}$ y reemplazamos para cada caso

$$p_i = \frac{p_i + p_l}{2}$$

antes de retomar la iteración correspondiente.

Utilizando el algoritmo anterior encontramos el valor mínimo del negativo de las funciones correspondientes en la función de log-verosimilitud de la familia GH y por tanto ubicamos sus puntos máximos. En este punto vale la pena retomar la importancia de la invarianza de los parámetros de la variable aleatoria GIG cuando se forman combinaciones lineales de los elementos del vector aleatorio X .

El algoritmo anterior en conjunto con la elección sistemática del punto inicial nos llevan a una estimación de la matriz de covarianzas subyacente bajo la distribución GH y constituye una aproximación de la matriz de covarianzas poblacional, la cual permite reducir las contribuciones aleatorias que afectan la estimación de un portafolio en el escenario de Markowitz.

3.3 Una prueba empírica para distribuciones multivariadas continuas

En general, el tema de la bondad de ajuste para distribuciones de probabilidad multivariada resulta escaso. Existen amplios estudios para el caso normal multivariante, pero poco se ha dicho del caso general. Si bien es cierto que algunos autores proponen extender la prueba de Kolmogorov Smirnov al caso n dimensional, Loudin (2003) y Justel (1997), no existe aún una aplicación e implementación para dimensiones superiores a dos. Las propuestas no paramétricas mediante estimaciones de kernel multivariadas Gábor (2004) no pueden generalizarse de manera sencilla y adolecen por tanto de problemas de implementación numérica hacia dimensiones superiores.

En esta sección incorporamos un test sencillo y de reciente aparición en la literatura para distribuciones multivariadas continuas de cualquier dimensión, McAssey (2013). No obstante, en nuestro contexto de la familia GH requerimos añadir dos algoritmos com-

plementarios para implementar la prueba, ya que la simulación de una variable aleatoria GH multivariada requiere simular una variable aleatoria *GIG*, Dagpunar (1989) y Kinderman (1977), así como una variable aleatoria normal multivariada, Ross (2012), de manera alternada. Ambos procedimientos se complementan para simular una distribución de probabilidad GH.

1. Primero mediante el algoritmo de Dagpunar simulamos una muestra aleatoria de tamaño n de una población cuya distribución es *GIG* con parámetros λ, χ, ψ y la denotamos por $W_i \sim GIG(\lambda, \chi, \psi)$.

2. Dado que un vector aleatorio GH, $X \sim GH(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ cumple que

$$X|W \sim N_m(\mu + W\gamma, W\Sigma),$$

entonces simulamos una muestra aleatoria de tamaño n de vectores $X_i \sim GH(\lambda, \chi, \psi, \mu, \Sigma, \gamma)$ mediante la normal multivariante $X_i|W_i$.

La prueba de hipótesis de McAssey (2013) requiere intensivamente de la simulación de los puntos anteriores 1 y 2, pues gracias a eso valoramos satisfactoriamente la bondad de ajuste hacia la distribución hiperbólica generalizada multivariada y reforzamos la estimación e inferencia de nuestro algoritmo EM modificado.

Si adaptamos el algoritmo empírico de McAssey (2013) a nuestro contexto de la familia GH, tenemos el siguiente procedimiento:

a) Asumimos que la hipótesis nula es verdadera, es decir, asumimos que las series de tiempo involucradas siguen una distribución hiperbólica generalizada multivariada de dimensión n y estimamos mediante el algoritmo EM modificado los parámetros correspondientes.

b) Generamos una muestra aleatoria u_1, u_2, \dots, u_m de vectores aleatorios $n \times 1$ de

tamaño 10000 con base en la distribución de probabilidad GH estimada en el inciso a). El algoritmo de simulación se apoya en el procedimiento de Dagpunar (1989) para simular una variable aleatoria GIG y en el algoritmo Ross (2012) para simular una normal multivariante, ya que recordemos se cumplen las siguientes condiciones

$$X|W \sim N_n(\mu + w\beta\Delta, w\Delta)$$

$$W \sim GIG(\lambda, \gamma, \delta)$$

c) Calculamos la distancia de Mahalanobis para las series de la muestra simulada en el inciso b) mediante la media $\hat{\mu}$ y la varianza $\hat{\Delta}$

$$\hat{d}_i = \sqrt{(u_i - \hat{\mu})' \hat{\Delta}^{-1} (u_i - \hat{\mu})}.$$

d) Luego para $t \in (0, 2\text{máx}\{\hat{d}_i\})$ calculamos

$$\hat{G}_m(t) = \frac{1}{m} \sum_{i=1}^m I(\hat{d}_i \leq t)$$

donde I es una función indicadora apropiada.

e) Elegimos una partición $\{p_0, p_1, \dots, p_T\}$ en el intervalo $[0, 1]$ tal que $p_0 = 0, p_T = 1$, y calculamos $q_0 = 0$,

$$q_j = \text{mín} \left\{ t \in \mathbb{R} : \hat{G}_m(t) \geq p_j \right\}$$

para $j = 1, 2, \dots, T - 1$ y $q_T = \infty$.

f) Estimamos $E_j = n(p_j - p_{j-1})$ y contamos O_j observaciones \hat{d}_i en el intervalo

$(q_{j-1}, q_j]$ para $j = 1, 2, \dots, T - 1$ y $(q_{T-1}, \infty]$. De esta manera el estadístico de prueba es

$$A_T = \sum_{j=1}^T \frac{|E_j - O_j|}{E_j}$$

Dado que la distribución de A_T varía según el conjunto de datos que se analice, encontramos el valor p para rechazar o no rechazar la hipótesis nula por simulación. Repetimos todos los incisos anteriores para una muestra aleatoria de tamaño $M = 1000$, McAssey (2013), sobre la distribución de probabilidad GH ajustada y rechazamos H_0 si valor p es menor que algún nivel de significancia dado α ,

$$\text{valor } p = \frac{1}{M} \sum_{i=1}^M I(A_{T,i} > A_T).$$

Capítulo 4

Aplicaciones y ejemplos

En esta parte presentamos algunos ejemplos donde la estimación de la matriz de covarianzas bajo los procedimientos comentados anteriormente resulta útil. Las estimaciones que se muestran ilustran los métodos estadísticos tanto de matrices aleatorias como de la distribución hiperbólica generalizada. Gran parte de las estimaciones matriciales no se muestran completamente en este documento, aunque las partes relevantes se colocaron en los anexos al final de esta tesis, pues resultan útiles para visualizar los distintos resultados.

4.1 Aspectos numéricos para estimar la distribución GH

En esta sección se presenta una modificación al algoritmo EM de máxima verosimilitud para estimar la distribución de probabilidad hiperbólica generalizada. La estimación que se realiza considera libre al parámetro λ de la función de Bessel modificada de tercer orden y toma sistemáticamente una aproximación inicial para reducir el número de iteraciones. La bondad de ajuste se valora con la función de log-verosimilitud y mediante una prueba empírica para distribuciones multivariadas continuas, McAssey (2013).

La familia de distribuciones hiperbólicas generalizadas puede obtenerse mediante una mezcla ponderada de una distribución normal multivariada y una distribución inversa gaussiana generalizada (GIG). Esta especificación queda en términos de los parámetros λ , χ y ψ de la GIG, los parámetros de la normal multivariante μ, Σ , y un vector de sesgo γ .

Existen diferentes parametrizaciones para la distribución GH, en este trabajo consideramos la especificación presentada por Protasov (2004), donde se resuelve el problema de especificación colocando el determinante de la matriz Δ igual a uno. En ese caso y bajo la definición presentada por McNeil, Frey y Embrechts (2005), se tiene la gran ventaja de que los parámetros de la GIG asociada (λ, χ, ψ) en la mezcla media-varianza son invariantes. Este resultado nos permite obtener un punto inicial sistemático para la ejecución del algoritmo EM en el cálculo de los parámetros de la distribución multivariada GH. Dada la invarianza de los parámetros (λ, χ, ψ) , empleamos una combinación lineal cuyo resultado implica una distribución hiperbólica univariada y obtenemos mediante máxima verosimilitud directa los valores óptimos $(\lambda_0, \chi_0, \psi_0)$. Esta estimación la tomamos como punto inicial para el algoritmo EM y procedemos mediante Protasov (2004) con λ fijo.

De esta forma, empleamos las series de tipo de cambio del franco suizo, el marco alemán, la libra esterlina, el dólar canadiense y el yen japonés, en relación al dólar americano, para el periodo que comprende del 07/02/1986 al 10/31/2001. Este conjunto de datos proveniente de Bloomberg es la muestra empleada por Protasov (2004). El objetivo es comparar las estimaciones considerando el parámetro fijo en $\lambda = -1/2$ versus la estimación con el parámetro λ libre y el punto inicial sistemático.

Cuadro 4.1. Estimación con parámetro λ fijo.

Estimación mediante el algoritmo EM con $\lambda=-0.5$					
Iteraciones	1071				
Valor-p	0.112				
Log-verosimilitud	79948.62				
$\hat{\lambda}$	-0.500				
$\hat{\delta}$	0.005				
$\hat{\gamma}$	272.434				
Matriz Δ					
Moneda	CHF	DEM	GBP	CAD	JPY
CHF	2.84577	2.37786	1.51159	0.00970	1.26953
DEM	2.37786	2.34835	1.40543	0.02615	1.12566
GBP	1.51159	1.40543	1.96464	0.03819	0.78977
CAD	0.00970	0.02615	0.03819	0.46850	-0.01940
JPY	1.26953	1.12566	0.78977	-0.01940	2.47008
Moneda	CHF	DEM	GBP	CAD	JPY
$\hat{\beta}$	36.3603	-25.7984	-17.2824	-9.0025	10.4371
$\hat{\mu}$	-0.0005	-0.0003	0.0001	0.0001	-0.0005

En el cuadro 4.1 se presentan las estimaciones considerando el parámetro fijo en $\lambda = -1/2$ y tomando el punto inicial $\tilde{\gamma} = 1$, $\tilde{\mu} = 0$, $\tilde{\beta} = 0$ y $\tilde{\Sigma} = I$ (matriz identidad). En contraste, en el cuadro 4.2 se muestran los resultados tomando en cuenta el punto inicial sistemático que resulta de maximizar la función de verosimilitud univariada para la combinación lineal $Y = B_j X$ sobre la base canónica de \mathbb{R}^5 y considerando el algoritmo EM con el parámetro λ libre. En concreto, nuestro punto inicial es $\lambda = -0.97$, $\tilde{\gamma} = 1.31$, $\tilde{\mu} = 0$, $\tilde{\beta} = 0$ y $\tilde{\Sigma} = I$ (matriz identidad). Observemos que el valor de log-verosimilitud

es mayor cuando realizamos la estimación con λ no fijo y que el valor p de la prueba empírica de bondad de ajuste es superior también en este caso. Sin mencionar que el número de iteraciones es más pequeño y evitamos complicaciones de estabilidad a lo largo de las iteraciones.

Cuadro 4.2. Estimación con parámetro λ no fijo.

Estimación mediante el algoritmo EM con λ libre y punto inicial sistemático					
Iteraciones	353				
Valor-p	0.437				
Log-verosimilitud	79952.42				
λ	-0.327				
δ	0.005				
γ	264.299				
Matriz Δ					
Moneda	CHF	DEM	GBP	CAD	JPY
CHF	2.72143	2.14245	1.42368	0.00846	1.32138
DEM	2.14245	2.16606	1.28088	0.03283	1.12272
GBP	1.42368	1.28088	1.86350	0.03869	0.77683
CAD	0.00846	0.03283	0.03869	0.43249	0.00932
JPY	1.32138	1.12272	0.77683	0.00932	2.33139
Moneda	CHF	DEM	GBP	CAD	JPY
β	22.4104	-15.9031	-13.7836	-10.0267	12.5831
μ	-0.0005	-0.0002	0.0001	0.0001	-0.0006

Las estimaciones anteriores muestran como la elección de un punto inicial sistemático para la implementación del algoritmo EM permite reducir el número de iteraciones en la estimación de los parámetros de la familia GH para el caso de Protasov (2004). Al

mismo tiempo el ajuste libre del parámetro λ permite alcanzar un valor óptimo superior en la función de log-verosimilitud de la familia GH.

Los resultados anteriores se corroboran mediante un test empírico de bondad de ajuste multivariado y ratifican con un valor p significativo que la distribución multivariada estimada con un λ libre se ajusta en mejor medida a las series de los rendimientos que la distribución con $\lambda = -0.5$. De esa manera la familia GH se constituye como una herramienta robusta para ajustar y simular diferentes series de tiempo no normales.

4.2 Variación del rendimiento de diferentes commodities durante el periodo 2000-2012

En esta sección aplicamos la teoría de matrices aleatorias (TMA) para analizar la variación del rendimiento de diferentes commodities para el periodo que va de enero del año 2000 al mes de diciembre de 2012. Se utiliza la matriz de correlaciones de los rendimientos diarios de una muestra de ochenta y un activos, donde se incluyen acciones representativas del NASDAQ, NYSE y diversos commodities. Mediante los valores propios de la matriz de correlaciones se filtra la variación libre de ruido de los rendimientos, y se encuentra evidencia estadística para señalar que la participación de algunos commodities en la variación sistemática común del total de activos se incrementó después de la crisis de 2008-2009.

La matriz de correlación empírica de un conjunto de rendimientos para una serie de activos es una herramienta relevante en finanzas. Recordemos simplemente que la matriz de correlaciones es el punto de partida en la teoría de selección de portafolios. En particular, si tenemos N activos con precio p_i en el tiempo t , $i = 1, 2, \dots, N$, $t = 0, 1, \dots, T$,

entonces el rendimiento geométrico se define como

$$r_{it} = \ln(p_{it}) - \ln(p_{i,t-1}).$$

La matriz empírica de correlaciones contiene contribuciones aleatorias, Laloux, Cizeau y Bouchaud (1999) y por ello resulta apropiado analizar las propiedades de la matriz de correlación empírica bajo la hipótesis nula de una matriz de correlación aleatoria. El objetivo es distinguir en la matriz de correlación empírica la variación de los rendimientos atribuible a “ruido” y la parte que contiene efectivamente “información”.

La matriz de correlación que consideramos se construye a partir de los rendimientos diarios de una muestra con ochenta y un activos, donde figuran acciones de NASDAQ, NYSE y diferentes commodities. El periodo muestral abarca desde el 3 de enero de 2000 hasta el 31 de diciembre de 2012. Durante este periodo de tiempo resulta incluíble la crisis internacional situada en 2008-2009, donde la disminución de la demanda agregada global implicó una caída generalizada en los precios de los commodities a partir del segundo semestre de 2008, véase figura 4.1.

El aumento sostenido de los precios de los commodities durante los años 2003-2008 mostró una caída durante la crisis global de 2008. Existen varios factores que afectan los precios de los commodities, tales como la oferta y demanda, los mercados especulativos, las políticas monetarias, las cuestiones climáticas y la problemática geopolítica, Gern (2008).

Las cuestiones climáticas y geopolíticas son factores de corto plazo, así que no podrían explicar completamente la tendencia alcista desde 2009. Por el lado de la demanda, el crecimiento de las economías emergentes, debido a su baja eficiencia en el uso de la energía y una alta tasa en el uso de insumos intermedios en la producción, inciden directamente en

el aumento del consumo agregado de energía y en los precios de los insumos intermedios, Silvennoinen (2010).

En cualquier caso, es bien conocido que existe una mayor relación entre los mercados de commodities y los mercados de acciones y bonos, pues durante la década pasada se observó un aumento sustancial en la actividad de los mercados de futuros financieros asociados a commodities, así como una creciente participación de inversores institucionales, particularmente en los metales preciosos, Silvennoinen (2010), donde funcionaron como activos refugio ante la incertidumbre de los hechos que sobrevinieron a partir de 2008.

Esta afirmación práctica que aparece en diferentes medios requiere evidencia estadística sólida. En este documento inferimos si esta aseveración es correcta, y para ello analizamos la participación de la variación de los rendimientos de los commodities en la variación total de los rendimientos de una muestra aleatoria de activos. Para ello discernimos la variación atribuible a ruido mediante los valores y vectores propios de la matriz de correlaciones empíricas, Mehta (2004).

En primer lugar, estimamos la matriz de correlación empírica para los distintos rendimientos en cada uno de los dos periodos que se han considerado. Dadas las limitaciones de espacio que tenemos en este documento, no presentamos la matriz de correlación con todos sus elementos estimados. No obstante, para fines comparativos reportamos la diagonal de la matriz de covarianzas en el anexo 1. De esa forma podemos observar las diferencias que existen en las varianzas de los rendimientos de los activos antes y después de filtrar la matriz de covarianzas.

Los valores propios que contienen información genuina se ubican fuera del intervalo $[\lambda_{\min}, \lambda_{\max}]$ y se presentan para cada periodo en el anexo 2. Similarmente, los vectores propios correspondientes a la variación sistemática se muestran en los anexos 3 y 4.

En cada caso, aplicamos la transformación respectiva para hallar los valores propios

desplegados, Laloux (1999), y realizamos las pruebas de hipótesis para verificar si se cumplen las propiedades universales de TMA.

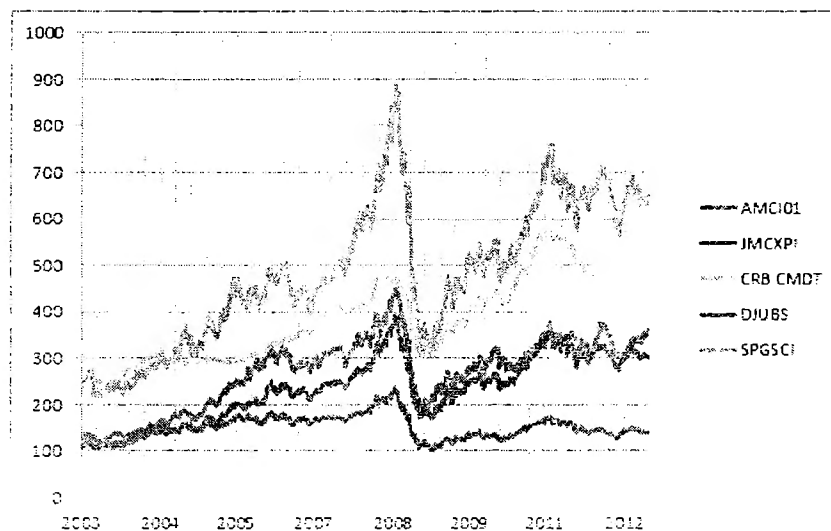


Figura 4.1. Variación de algunos índices representativos sobre commodities.

Fuente. Bloomberg.

En el cuadro 4.3 presentamos los resultados de los tests estadísticos para el primer y segundo vecino más cercanos de los valores propios desplegados. En cada caso, se puede observar que no podemos rechazar la hipótesis nula. La distribución de probabilidad empírica y la distribución de probabilidad teórica para la diferencia del primer y segundo vecino más cercano se ajustan razonablemente.

En otras palabras, existe evidencia para justificar que los valores propios que se en-

cuentran fuera del intervalo capturan información libre de ruido de la variación de los rendimientos de los activos.

En la figura 4.4 presentamos la distribución de probabilidad correspondiente a la diferencia de los valores propios desplegados, que corresponden al primer y segundo vecino más cercano, para el periodo que va de enero de 2000 al 15 de junio de 2008. Para el caso del periodo 2008-2012 se muestran las gráficas en la figura 4.3.

Por otra parte, cuando consideramos la varianza número se puede observar que conforme aumenta el valor de la longitud del intervalo que contiene a los valores propios desplegados, sucesivamente nos acercamos al logaritmo natural de la longitud del intervalo, véase figura 4. Esto apoya los resultados de las pruebas de hipótesis que se muestran en el cuadro 4.3.

Ahora bien, mediante el índice de participación del vector propio correspondiente al valor propio máximo podemos medir cuál es la participación que tienen los commodities en la variación total del rendimiento para los periodos que van de enero de 2000 a junio de 2008 y de junio de 2008 a diciembre de 2012.

El valor propio máximo para el primer periodo tiene un valor de 11.648 e implica 52 activos representativos que participan en la variación sistemática del rendimiento, mientras que para el periodo 2008-2012 el valor propio máximo 27.963 conlleva 64 activos representativos en la variación sistemática común, véanse anexos 3 y 4.

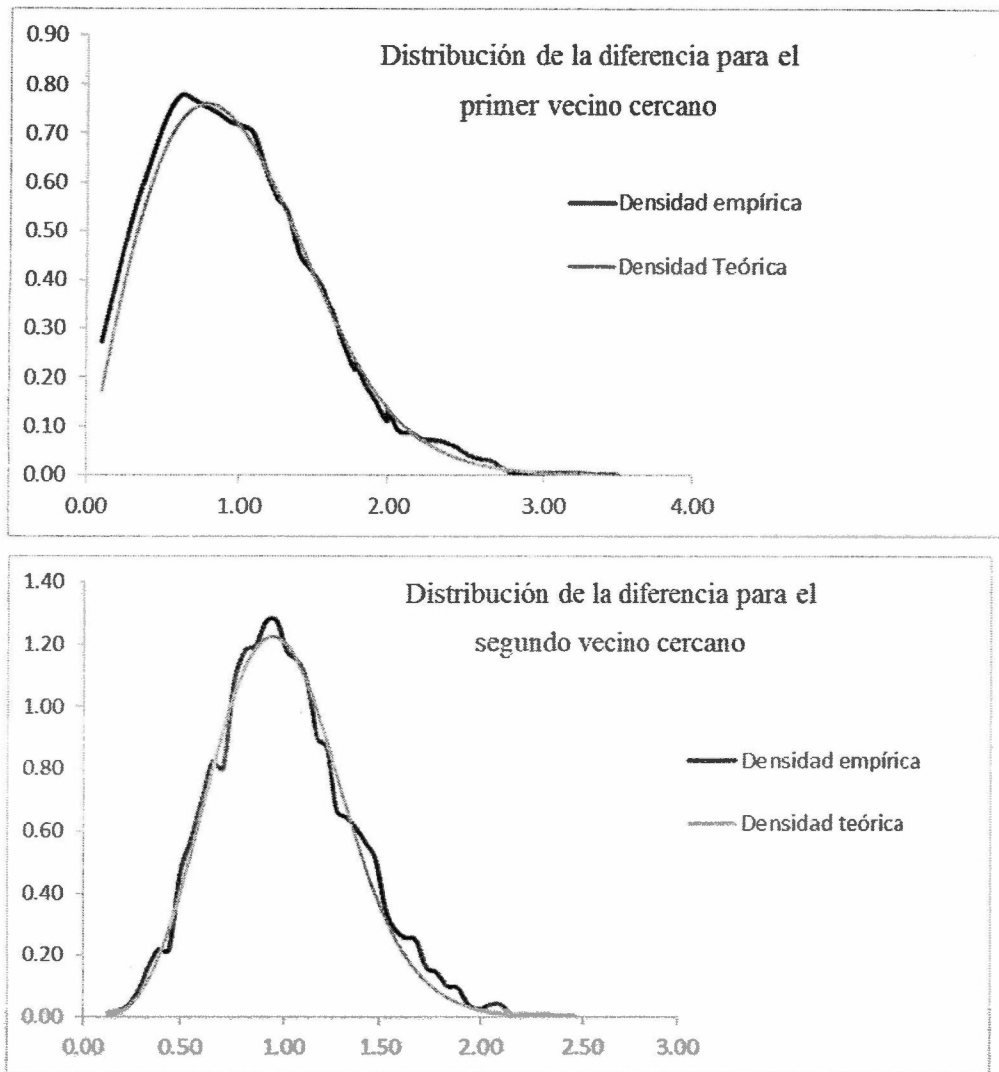


Figura 4.2. Distribuciones de probabilidad para la diferencia de los vecinos más cercanos en los valores propios desplegados, 2000-2008.

Cuadro 4.3. Pruebas de hipótesis para el periodo 2000-2012.

Periodo	Intervalo de los valores propios $[\lambda_{\text{mín}}, \lambda_{\text{máx}}]$	Valor p (Test de Kolmogorov)	
		Prueba primer vecino más cercano	Prueba segundo vecino más cercano
2000-2007	[0.646,1.431]	0.896	0.412
2008-2012	[0.561,1.566]	0.170	0.120

El cuadro 4.4 muestra que existe un incremento de 4.51% de participación de los rendimientos de los commodities (metales preciosos), en la variación sistemática del rendimiento de los activos. Es decir, se tiene evidencia para inferir que poblacionalmente los rendimientos de los commodities (metales preciosos) aumentaron su participación en la variación sistemática del conjunto total de activos de la muestra. Este hallazgo muestra indirectamente que los commodities (metales preciosos) atrajeron mayor variación del mercado en el segundo periodo que va de junio de 2008 a diciembre de 2013.

En contraste, bajo la distribución de probabilidad multivariada encontramos un incremento de 4.4% de participación de los rendimientos de los metales preciosos, en la variación sistemática del rendimiento de los activos. Este resultado se basa en una bondad de ajuste de la distribución de probabilidad GH razonable, pues el valor p asociado para la prueba de McAssey es 0.145 para el periodo que va de 2000 a 2008 y de 0.218 para el periodo que va de 2008 al año 2012. Las estimaciones de los parámetros GH tanto como de los vectores y valores propios se encuentran en los anexos, al final de este documento.

Cuadro 4.4. Participación de los commodities (metales preciosos) según TMA en la variación sistemática común del rendimiento 2000-2012.

Periodo	Participación en la variación sistemática del rendimiento de los activos totales de la muestra NASDAQ, NYSE y commodities	
	2000-2007	2008-2012
Acciones (NASDAQ/NYSE)	97.30%	92.79%
Commodities (metales preciosos)	2.70%	7.21%

Cuadro 4.5. Participación de los commodities según GH (metales preciosos) en sistemática común del rendimiento 2000-2012.

Periodo	Participación en la variación sistemática total del rendimiento de los activos totales de la muestra NASDAQ, NYSE y commodities	
	2000-2007	2008-2012
Acciones (NYSE/NASDAQ)	95.8%	91.4%
Commodities (metales preciosos)	4.2%	8.6%

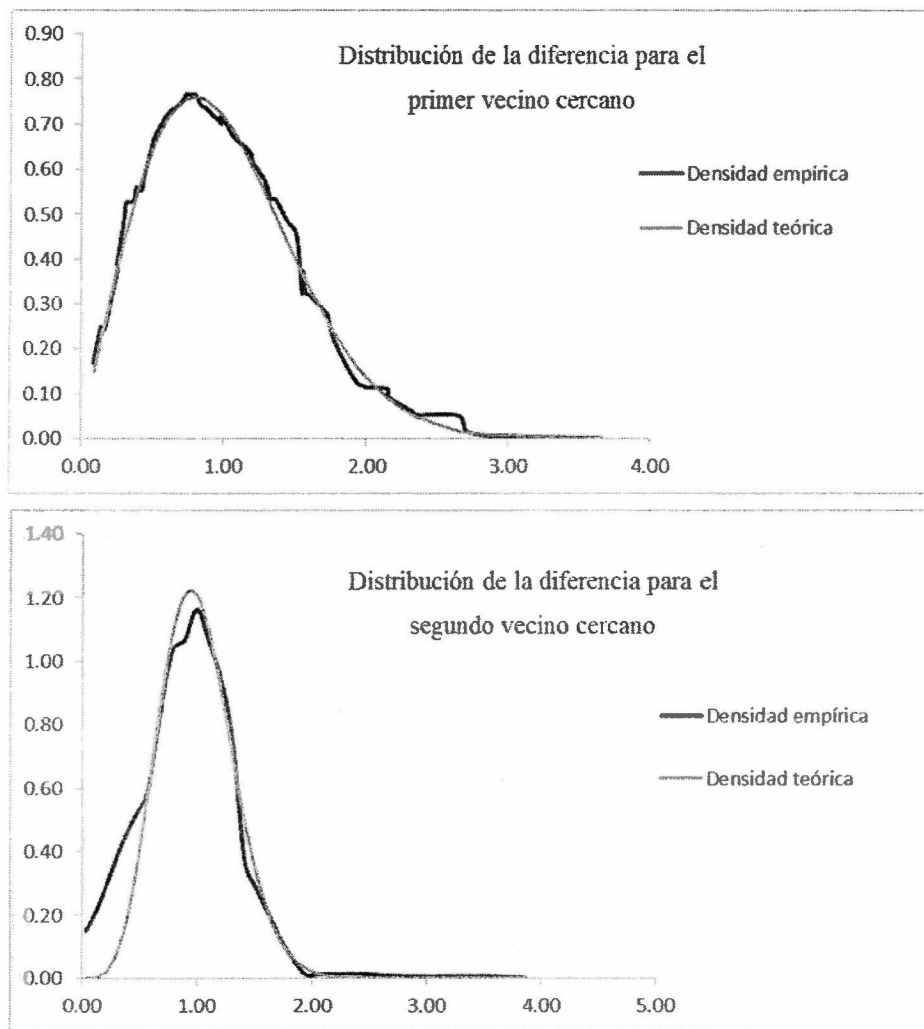


Figura 4.3. Distribuciones de probabilidad para la diferencia de los vecinos más cercanos en los valores propios desplegados, 2008-2012.

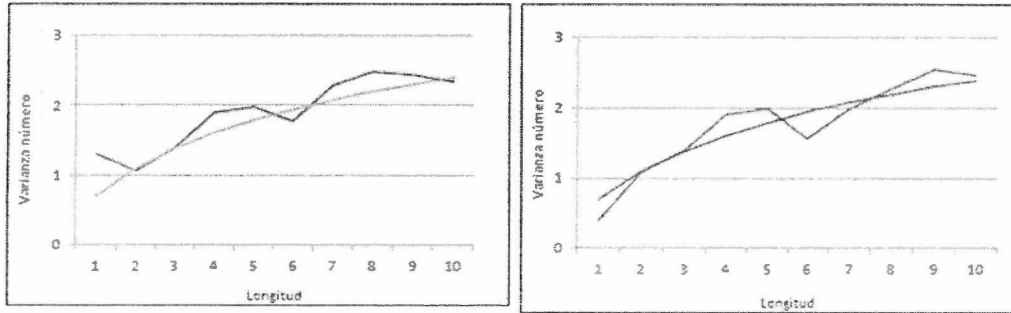


Figura 4.4. Comparación entre la varianza número y el logaritmo natural de la longitud para el periodo 2000-2008 (izquierda) y 2008-2012 (derecha).

Notemos que bajo el marco de la teoría de matrices aleatorias, encontramos evidencia para discernir la variación del rendimiento libre de ruido de algunos commodities en una muestra aleatoria de acciones de NASDAQ y NYSE.

Mediante el análisis de dos periodos de tiempo durante los años que van de 2000 a 2012, se verifican las pruebas de hipótesis que sustentan las propiedades universales de las matrices aleatorias. Los valores propios desplegados interiores de la matriz empírica de correlaciones de los rendimientos siguen las funciones de densidad de probabilidad predichas por la TMA, así que es posible distinguir los valores y vectores propios correspondientes para clasificar la información genuina de aquella producida por ruido a lo largo del periodo de la muestra.

Mediante el cociente inverso de participación y los vectores propios correspondientes a los valores propios máximos, se calcula la participación del grupo de commodities de metales preciosos en la variación sistemática común del total de los activos de la muestra, y se encuentra evidencia estadística para inferir que poblacionalmente los rendimientos de los commodities (metales preciosos) aumentaron su participación en la variación sis-

temática del conjunto total de activos de la muestra. En otras palabras, se encuentra indirectamente que los metales preciosos atrajeron mayor variación sistemática del mercado durante el periodo que va de junio de 2008 a diciembre de 2012. Es decir, existe evidencia para afirmar que efectivamente los metales preciosos funcionaron como activos refugio en el periodo posterior al año 2008, pues se incrementa su participación sistemática en la variación total.

4.3 Herramientas complementarias para construir portafolios eficientes

El trabajo desarrollado por Markowitz en los años cincuenta del siglo XX puede verse como el origen de la teoría moderna de portafolios. Markowitz asume un inversionista averso al riesgo que dispone de un capital inicial y que desea invertir en un conjunto de activos financieros con rendimientos futuros. En ese sentido, busca determinar la proporción de su ingreso a invertir en cada uno de estos activos para maximizar el rendimiento esperado del portafolio. El portafolio será eficiente siempre que el rendimiento esperado del mismo sea máximo comparado con otros portafolios posibles dado un nivel de riesgo dado. En la optimización de portafolios, la desviación estándar fue propuesta por Markowitz como una medida de riesgo.

Si suponemos que R_{ij} denota el rendimiento en el periodo j del activo i definimos

$$\mu_i = E[R_{ij}]$$

$$\sigma_i^2 = V[R_{ij}]$$

$$\sigma_{ik} = Cov[R_{ij}, R_{kj}]$$

para $i, k \in \{1, 2, \dots, n\}$, $i \neq k$ y $j \in \{1, 2, \dots, T\}$.

Si w_i denota el porcentaje de riqueza invertido en el activo i tal que

$$\sum_{i=1}^n w_i = 1,$$

el valor esperado μ_P y la varianza σ_P^2 del portafolio son

$$\begin{aligned}\mu_P &= \sum_{i=1}^n w_i \mu_i \\ \sigma_P^2 &= \sum_{i=1}^n w_i^2 \sigma_i^2 + \sum_{i < k} \sum_{i < k} w_i w_k \sigma_{ik}\end{aligned}$$

El problema de optimización a resolver para encontrar el portafolio de mínima varianza sin considerar un activo libre de riesgo es

$$\min_w \sum_{i=1}^n w_i^2 \sigma_i^2 + \sum_{i < k} \sum_{i < k} w_i w_k \sigma_{ik}$$

sujeto a la restricción $\sum_{i=1}^n w_i = 1$. Si empleamos notación matricial podemos reescribir el problema anterior como

$$\begin{aligned}\min_w w' \Sigma w \\ \text{s.a. } w' u = 1\end{aligned}$$

siendo u un vector de 1's. El lagrangiano asociado es

$$L = w' \Sigma w - \lambda [w' u - 1]$$

cuyas condiciones de primer orden son

$$\begin{aligned}\frac{\partial L}{\partial w} &= 2\Sigma w + \lambda u = 0 \\ \frac{\partial L}{\partial \lambda} &= w'u - 1 = 0\end{aligned}$$

Se puede despejar w de la primera ecuación

$$w = -\frac{1}{2}\lambda\Sigma^{-1}u$$

luego

$$1 = u'w = -\frac{1}{2}\lambda u'\Sigma^{-1}u$$

implica

$$w = \frac{\Sigma^{-1}u}{u'\Sigma^{-1}u}$$

El rendimiento esperado óptimo y la varianza mínima para este caso son

$$\begin{aligned}\mu_{P,w} &= w'\mu \\ \sigma_{P,w}^2 &= w'\Sigma w\end{aligned}$$

El conjunto de oportunidades de inversión es el conjunto de rendimientos esperados del portafolio, μ_P , y la desviación estándar del portafolio, σ_P , para todos los portafolios factibles tales que la suma de las ponderaciones, w_i , sumen la unidad. Podemos simplificar el análisis de este conjunto si nos enfocamos en los portafolios eficientes. Estos portafolios se encuentran en la frontera del conjunto de oportunidades de inversión por arriba del portafolio de varianza mínima. Este es el marco original de Markowitz, (1959), ganador del premio Nobel de Economía en 1990. Bajo este enfoque los inversionistas buscan los

portafolios que maximizan el rendimiento esperado dado un nivel de riesgo, siendo este último medido por la varianza del portafolio,

$$\begin{aligned} \text{máx } \mu_P &= w' \mu \\ \text{s.a } w' u &= 1 \\ \sigma_{P,0}^2 &= w' \Sigma w \end{aligned}$$

Este problema puede plantearse en forma dual, de forma que se minimiza el riesgo del portafolio para un nivel dado de rendimiento,

$$\begin{aligned} \text{mín } \sigma_P^2 &= w' \Sigma w \\ \text{s.a } w' u &= 1 \\ \mu_{P,0} &= w' \mu \end{aligned}$$

La frontera del conjunto de portafolios eficientes es una gráfica de μ_P versus σ_P para cada uno de los valores factibles $\mu_{P,0}$ de las soluciones del problema planteado arriba. El lagrangiano asociado para solucionar el problema de minimización es

$$L = w' \Sigma w + \lambda_1 [w' u - 1] + \lambda_2 [w' u - 1]$$

y las condiciones de primer orden

$$\begin{aligned} \frac{\partial L}{\partial w} &= 2 \Sigma w + \lambda_1 \mu + \lambda_2 u = 0 \\ \frac{\partial L}{\partial \lambda_1} &= w' \mu - \mu_{P,0} = 0 \\ \frac{\partial L}{\partial \lambda_2} &= w' u - 1 = 0 \end{aligned}$$

Si despejamos w de la primera ecuación

$$w = -\frac{1}{2}\lambda_1\Sigma^{-1}\mu - \frac{1}{2}\lambda_1\Sigma^{-1}u$$

y por las otras dos ecuaciones encontramos

$$\begin{aligned}\mu_{P,0} &= \mu'w = -\frac{1}{2}\lambda_1\mu'\Sigma^{-1}\mu - \frac{1}{2}\lambda_1\mu'\Sigma^{-1}u \\ 1 &= u'w = -\frac{1}{2}\lambda_1u'\Sigma^{-1}\mu - \frac{1}{2}\lambda_1u'\Sigma^{-1}u\end{aligned}$$

Esto implica un sistema de ecuaciones lineales para λ_1 y λ_2 ,

$$-\frac{1}{2}\begin{bmatrix} A & B \\ B & C \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} \mu_{P,0} \\ 1 \end{bmatrix}$$

donde

$$A = \mu'\Sigma^{-1}\mu$$

$$B = \mu'\Sigma^{-1}u$$

$$C = u'\Sigma^{-1}u$$

implican la solución

$$\begin{aligned} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} &= -2 \begin{bmatrix} A & B \\ B & C \end{bmatrix}^{-1} \begin{bmatrix} \mu_{P,0} \\ 1 \end{bmatrix} \\ \begin{bmatrix} A & B \\ B & C \end{bmatrix}^{-1} &= \frac{1}{AC - B^2} \begin{bmatrix} C & -B \\ -B & A \end{bmatrix} \\ \lambda_1 &= -\frac{2}{AC - B^2} (C\mu_{P,0} - B) \\ \lambda_2 &= -\frac{2}{AC - B^2} (A - B\mu_{P,0}) \end{aligned}$$

que lleva a la solución

$$w = \frac{C\mu_{P,0} - B}{AC - B^2} \Sigma^{-1} \mu + \frac{A - B\mu_{P,0}}{AC - B^2} \Sigma^{-1} u$$

La aleatoriedad presente en la matriz de correlaciones empírica tiene implicaciones importantes en la construcción de un portafolio óptimo bajo el esquema de Markowitz. Si consideramos que el rendimiento de un portafolio en el tiempo t está dado por

$$r(t) = \sum_{i=1}^N w_i r_{ti}$$

donde r_{ti} es el rendimiento del activo i en el tiempo t y w_i es la fracción de riqueza invertida en el activo i . El riesgo del portafolio puede cuantificarse por medio de la desviación estándar del activo i tal que la varianza es

$$\sigma_i^2 = \sum_{i=1}^N \sum_{j=1}^N w_i w_j C_{ij} \sigma_i \sigma_j$$

donde σ_i es la volatilidad promedio de r_{ti} , y C_{ij} son los elementos de la matriz de cor-

relación de rendimientos.

El portafolio óptimo puede encontrarse si minimizamos σ^2 sujeto a las restricciones

$$\begin{aligned} \sum_{i=1}^N w_i &= 1 \\ w_1, w_2, \dots, w_N &\geq 0 \\ \sum_{i=1}^N w_i r_{ti} &= \mu_0 \end{aligned}$$

mediante programación cuadrática, Engels (2004).

Los pesos w_i determinan el portafolio óptimo y dependen sustancialmente de la matriz de correlación C . En particular, si existen componentes aleatorios relevantes en los elementos estimados de esta matriz, el efecto del ruido es clave. Por ello estimamos la matriz de correlaciones mediante la teoría de matrices aleatorias y bajo la distribución de probabilidad GH para el periodo completo 2000-2012, véase la sección de anexos. En el cuadro 4.6 se presentan los valores p asociados a las pruebas de hipótesis de TMA y en el cuadro 4.7 el valor p correspondiente a la bondad de ajuste para la distribución GH, en cada caso no se rechaza la hipótesis nula con un nivel de significancia de $\alpha = 0.05$.

Cuadro 4.6. Pruebas de hipótesis de TMA para el periodo 2000-2012.

Periodo	Intervalo de los valores propios [$\lambda_{\min}, \lambda_{\max}$]	Valor p (Test de Kolmogorov)	
		Prueba primer vecino más cercano	Prueba segundo vecino más cercano
2000-2012	[0.646, 1.431]	0.496	0.312

Cuadro 4.7. Bondad de ajuste de la distribución GH sobre los periodos 2000-2008, 2008-2012 y 2000-2012.

Periodo	Prueba de McAssey	
	Estadístico de prueba	Valor p
2000-2008	105.06	0.345
2008-2012	110.06	0.231
2000-2012	100.58	0.465

En el cuadro 4.8 se presentan los portafolios de mínima varianza para el periodo 2000-2012 mediante la matriz de correlación de rendimientos empírica, la matriz de correlación filtrada mediante TMA y la matriz de correlación según la distribución de probabilidad GH. Bajo estos tres casos encontramos diferentes portafolios óptimos y por ende, distintos rendimientos esperados con varianzas diferentes. Para comparar las tres carteras obtenidas empleamos el coeficiente de variación, ya que este parámetro permite comparar dos grupos de datos que tienen distinta media.

Los resultados mostrados en el cuadro 4.8 señalan que el menor coeficiente de variación se obtiene bajo la matriz de correlación estimada bajo GH, lo cual indica indirectamente que la matriz de correlaciones estimada se encuentra más cerca de la matriz de correla-

ciones poblacional asociada al conjunto de datos 2000-2012.

Cuadro 4.8. Portafolio de Mínima Varianza.

Portafolio de Mínima Varianza			
	Markowitz	TMA	GH
Rendimiento medio	0.0002150	0.0002158	0.0002141
Varianza	0.0000235	0.0000265	0.0000410
Desviación estándar	0.0048460	0.0051452	0.0064028
Coefficiente de variación	0.0443637	0.0419439	0.0334461

Mediante la prueba de bondad de ajuste de McAssey podemos afirmar que la distribución multivariada para el conjunto de rendimientos sigue una función de la familia GH, entonces el portafolio de mínima varianza encontrado bajo GH nos dice que la cartera encontrada de manera tradicional por el procedimiento de Markowitz subestima la varianza poblacional del portafolio. En el cuadro 4.9, se muestra la varianza del portafolio de Markowitz cuando se valora bajo la matriz de covarianzas de TMA y GH.

Cuadro 4.9 Portafolio de Mínima Varianza.

Portafolio de Mínima Varianza		
	GH	TMA
Rendimiento medio	0.0002150	0.0002150
Varianza	0.0000473	0.0000465
Desviación estándar	0.0068772	0.0068171
Coefficiente de variación	0.0312608	0.0315362

Gráficamente tenemos la siguiente figura, la cual ilustra como el procedimiento tradicional de Markowitz subestima la varianza poblacional del portafolio .

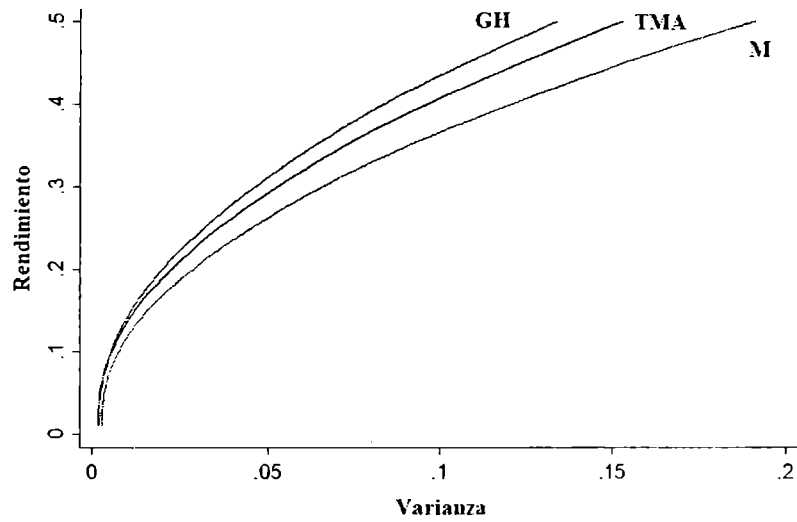


Figura 4.5. Portafolio de Markowitz.

Conclusiones

En esta tesis se ratifica la importancia de la matriz de covarianza empírica para el análisis de los rendimientos en una serie de activos. En particular, abordamos dos diferentes posibilidades para distinguir la "señal de ruido", y así filtrar la matriz de covarianzas que contiene información verdadera para el control de riesgos: teoría de matrices aleatorias y distribución hiperbólica generalizada.

Las pruebas de hipótesis bajo TMA permiten discernir la variación por ruido. No obstante, no añaden información sobre la distribución de probabilidad implícita, por ello se comparan sus resultados con la estimación paramétrica de la distribución GH. La calibración de la familia GH se realiza mediante algunos complementos numéricos al algoritmo EM, en particular se elige de manera sistemática un punto inicial para empezar las iteraciones y dada la sensibilidad de la función de Bessel, ante la variación numérica del parámetro λ , se emplean conjuntamente procedimientos numéricos que no hacen uso de derivadas con las propiedades de invarianza de los parámetros de la GIG.

Los resultados de las estimaciones son razonables y extienden la aplicación de la familia GH hacia otras aplicaciones, pues el valor de log-verosimilitud que se alcanza es superior al que se alcanzaría bajo el supuesto de λ fijo. Sin mencionar que el valor p asociado a la prueba de bondad de ajuste multivariado permite no rechazar la hipótesis nula de GH con un nivel de significancia superior a 10%. Estos resultados nos señalan

conjuntamente que la matriz de covarianzas estimada bajo la familia GH aproxima de manera robusta a la matriz de varianza-covarianza poblacional.

De esta manera, el procedimiento de estimación considera simultáneamente la bondad de ajuste de los parámetros calculados y facilita la aplicación de la distribución hiperbólica GH hacia la construcción de portafolios. Este ajuste permite valorar el impacto que tiene la matriz de covarianzas muestral dentro de los portafolios de Markowitz e inferir el efecto que tienen los choques aleatorios de los datos muestrales en la matriz de correlación empírica.

Específicamente, el portafolio de mínima varianza señalado por Markowitz alcanza una varianza menor si empleamos las técnicas TMA o GH, donde GH permite construir un portafolio cuyo riesgo es el más pequeño, pues la matriz de covarianzas utilizada bajo GH es sumamente cercana a la matriz de covarianzas poblacional de los rendimientos. Esta afirmación puede verificarse mediante la prueba de McAssey (2013) y permite afirmar que la filtración por ruido permite refinar la construcción de portafolios de Markowitz.

Adicionalmente, el proceso de filtración permite eliminar el componente ruidoso en periodos de alta volatilidad, lo cual conduce a estimaciones razonables de la variación sistemática de los rendimientos en la toma de decisiones.

Referencias

Abramowitz, M. y Stegun, I. A. (1972). Handbook of Mathematical Functions. New York, USA: Dover; Chapter 9 Bessel Functions of Integer Order.

Bai, Z.D. (1999). Methodologies in Spectral Analysis of Large Dimensional Random Matrices: A review. *Statistica Sinica*, 9(3), 611-677.

Barndorff-Nielsen, O.E. (1977). Exponentially decreasing distributions for the logarithm of the particle size. *Proceedings of the Royal Society. London. Series A. Mathematical and Physical Sciences*, 353:401-419.

Blæsild, P. and Sørensen, M.(1992). HYP: a computer program for analyzing data by means of the hyperbolic distribution. *Research Report 248*, Department of Theoretical Statistics, University of Aarhus.

Bouchaud, J.P. y Potters, M. (2003). Theory of Financial Risks: From Statistical Physics to Risk Management. Cambridge, UK. Cambridge University Press.

Breymann (2011). GHyp: A package on generalized hyperbolic distributions. CRAN, 1-26.

Brody, T. A., Flores, J., French, J.B., Mello, P., Pandey, A. y Wong, S. (1981). Random Matrix Physics Spectrum and Strength Fluctuations. *Rev. Mod. Phys.*, 53, 3, 385-480.

Burda, Z. y Jurkiewicz, J. (2004). Signal and Noise in Financial Correlation Matrices. *Physica A, Statistical Mechanics and its Applications*, 344, 67-72.

Dagpunar, J. (1989). An Easily Implemented Generalised Inverse Gaussian Generator. *Communications in Statistics. Simulation and Computation* , 703-710.

Dempster, A.P., Laird, N.M. and Rubin, D.B.(1977). Maximum likelihood from incomplete data via the EM algorithm. *J. Roy. Statist. Soc., Ser B* 39 1-38.

Eberlein, Ernst and Keller, Ulrich (1995). Hyperbolic distributions in finance. *Bernoulli*, 1:281-299.

Engels, M. (2004). Portfolio Optimization: beyond Markowitz. Master´s Thesis. Universiteit Leiden.

Justel, A., Peña, D. and Zamar, R. (1997). A multivariate Kolmogorov-Smirnov test of goodness of fit. *Statistics and Probability Letters*, 251-259.

Giada, L. y Marsili, M. (2001). Data Clustering and Noise Undressing of Correlation Matrices. *Phys.Rev.E*, 63, 061101.

Gábor, J.S. and rizzo, M.L. (2004). Testing for equal distributions in high dimension. Bowling Green State University and Ohio University. Working paper.

Gupta, A.K. y Nagar, D. K. (2000). Matrix Variate Distributions. New York, USA. Chapman and Hall.

Kelley C. T. (1999). Detection and Remediation of Stagnation in the Nelder Mead Algorithm Using a Sufficient Decrease Condition. *SIAM J.*, USA.

Kinderman, A. J. (1977). Computer Generation of Random Variables Using the Ratio of Uniform Deviates. *ACM Transactions on Mathematical Software*, 257-260.

Laloux, L., Cizeau, P. y Bouchaud, J. P. (1999). Noise Dessing of Financial Correlation Matrices. *Phys. Rev. Lett.*, 83, 1467-1470.

Laloux, L., Cizeau, P., Potters, M. y Bouchaud, J. P. (2000). Random Matrix Theory and Financial Correlations. *International Journal of Theoretical and Applied Finance*, 3(3), 391-397.

Liu, C. and Rubin, D.(1995). ML estimation of the t distribution using EM and its extensions, ECM and ECME. *Statistica Sinica*, 5(1995), 19-39.

Loudin, J. and Hannu, M. (2003). A multivariate method for comparing n-dimensional distributions. Phystat 2003, SLAC, Stanford, California.

Lütli (2011). Package GHyp. CRAN, 1-50.

Markowitz, H. (1959). Portfolio Selection: Efficient Diversification of Investments. New York, USA: Wiley.

Maslov S. (2001). Measures of Globalization Based on Cross-Correlations of World Financial Indices. Physica A, 301, 397-406.

Mehta, L. M. (2004). Random Matrices. California, USA. Elsevier.

Meyer, H., Angles, J.C., Maillard, J.M. (1997). Random Matrix Theory and Classical Statistical Mechanics. *Phys. Rev. E*, 55, 380-392.

McAssey, M. (2013). An empirical goodness-of-fit test for multivariate distributions. *Journal of Applied Statistics*, 40,5,1120-1131.

McNeil, A., Frey, R. and Embrechts, P.(2005). Quantitative Risk Management: Concepts, Techniques and Tools. Princeton University Press.

Nelder, J. y Mead, R. (1965). A simplex method for function minimization. *The Computer Journal*, USA.

Paoella, M. S. (2007). *Intermediate Probability*. West Sussex, England. John Wiley & Sons.

Prause, K.(1999). The Generalized hyperbolic model: estimation, financial derivatives, and risk measures. Tese de Doutorado, University of Freiburg.McNeil et al. (2005).

Protasov, Rostislav(2004). EM-based maximum likelihood parameter estimation for multivariate generalized hyperbolic distributions with λ fixed. *Statistics and Computing*. Vol. 14, issue 1.

R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org/>.

Scott, D. (2012). The Generalized Hyperbolic Distribution. CRAN , 1-77.

Trejo Becerril, Bárbara Ruth; Lorenzo Valdés, Arturo; Núñez Mora, José Antonio (2006). Distribución de los rendimientos del mercado mexicano accionario. *Estudios Económicos*. México.

Wenbo Hu (2005). Calibration of multivariate generalized hyperbolic distributions using the EM algoritm, with applications in risk management, portfolio optimization and portfolio credit risk. Working paper. The Florida State University.

Anexo 1.A Acciones de NASDAQ, NYSE y commodities que se utilizaron en la muestra para el periodo 2000-2012. Los datos constan de precios de cierre diarios reportados por Yahoo Finance y Bloomberg.

No.	Acciones/Commodities		No.	Acciones/Commodities	
1	A	Agilent Technologies Inc. (A)-NYSE	30	CRMT	America's Car-Mart Inc. (CRMT) - NasdaqGS
2	AA	Alcoa, Inc. (AA)-NYSE	31	CRUS	Cirrus Logic Inc. (CRUS) - NasdaqGS
3	AAPL	Apple Inc. (AAPL) - NasdaqGS	32	CSTR	Coinstar, Inc. (CSTR) - NasdaqGS
4	ACEITE DE PALMA	KO1 Comdty	33	CWTR	Coldwater Creek Inc. (CWTR)-NasdaqGS
5	ADM	Archer Daniels Midland Company (ADM) -NYSE	34	D	Dominion Resources, Inc. (D)-NYSE
6	AF	Astoria Financial Corporation (AF)-NYSE	35	DECK	Deckers Outdoor Corp. (DECK) - NasdaqGS
7	AIG	American International Group, Inc. (AIG) -NYSE	36	DELL	Dell Inc. (DELL) - NasdaqGS
8	ALGODÓN	CT1 Comdty	37	DUK	Duke Energy Corporation (DUK) -NYSE
9	ALUMINIO	LMAHDY LME Comdty	38	EBAY	eBay Inc. (EBAY) - NasdaqGS
10	AMGN	Amgen Inc. (AMGN) - NasdaqGS	39	ELNK	EarthLink Inc. (ELNK) - NasdaqGS
11	AMZN	Amazon.com Inc. (AMZN) - NasdaqGS	40	EMN	Eastman Chemical Co. (EMN) -NYSE
12	ANNB	Annapolis Bancorp Inc. (ANNB)-NasdaqCM	41	EXPD	Expeditors International of Washington Inc. (EXPD)-NasdaqGS
13	ANSS	Ansys, Inc. (ANSS)-NasdaqGS	42	F	Ford Motor Co. (F) -NYSE
14	APC	Anadarko Petroleum Corporation (APC) -NYSE	43	FSYS	Fuel Systems Solutions, Inc. (FSYS) - NasdaqGS
15	ARMH	ARM Holdings plc (ARMH)-NasdaqGS	44	GANADO	LC1 COMB Comdty
16	ARROZ	RR1 COMB Comdty	45	GOMA	JN1 COMB Comdty
17	BAC	Bank of America Corporation (BAC)-NYSE	46	IBM	International Business Machines Corporation (IBM) -NYSE
18	BBY	Best Buy Co. Inc. (BBY)-NYSE	47	JNPR	Juniper Networks, Inc. (JNPR) -NYSE
19	BLK	BlackRock, Inc. (BLK) -NYSE	48	JUGO DE NARANJA	JO1 Comdty
20	BSX	Boston Scientific Corporation (BSX) -NYSE	49	L	Loews Corporation (L)-NYSE
21	C	Citigroup, Inc. (C)-NYSE	50	LSCC	Lattice Semiconductor Corporation (LSCC)-NasdaqGS
22	CACAO	CC1 Comdty	51	M	Macy's, Inc. (M)-NYSE
23	CELG	Celgene Corporation (CELG) - NasdaqGS	52	MADERA	LB1 COMB Comdty
24	CHK	Chesapeake Energy Corporation (CHK) -NYSE	53	MSFT	Microsoft Corporation (MSFT) - NasdaqGS
25	CIENA	CIENA Corp. (CIEN) - NasdaqGS	54	NVE	NV Energy, Inc. (NVE) -NYSE
26	CLF	Cliffs Natural Resources Inc. (CLF) -NYSE	55	ORCL	Oracle Corporation (ORCL) - NasdaqGS
27	CMC	Commercial Metals Company (CMC) -NYSE	56	ORO	XAU BGN Curncy
28	COBRE	LMCADY LME Comdty	57	PALADIO	XPD BGN Curncy
29	COST	Costco Wholesale Corporation (COST) - NasdaqGS	58	PFH	Cabco Tr Jcp 7.625 Common Stock (PFH)-NYSE

Anexo 1.B Acciones de NASDAQ, NYSE y commodities que se utilizaron en la muestra para el periodo 2000-2012. Los datos constan de precios de cierre diarios reportados por Yahoo Finance y Bloomberg.

No.	Acciones/Commodities		No.	Acciones/Commodities	
59	PLATA	XAG Silver Spot	71	USG	USG Corporation (USG)-NYSE
60	PLATINO	XPTXPD CMPN Curncy	72	USNA	USANA Health Sciences Inc. (USNA)-NYSE
61	PLOMO	LMPBDY LME Comdty	73	UTX	United Technologies Corp. (UTX)-NYSE
62	PNRA	Panera Bread Co. (PNRA)-NasdaqGS	74	VOD	Vodafone Group Public Limited Company (VOD) -NasdaqGS
63	R	Ryder System, Inc. (R)-NYSE	75	WDC	Western Digital Corporation (WDC) -NasdaqGS
64	RFMD	RF Micro Devices Inc. (RFMD) -NasdaqGS -NYSE	76	WLT	Walter Energy, Inc. (WLT) -NYSE
65	RIMM	Research In Motion Limited (RIMM) -NasdaqGS	77	X	United States Steel Corp. (X)-NYSE
66	SAN	Banco Santander, S.A. (SAN)-NYSE	78	XOM	Exxon Mobil Corporation (XOM)-NYSE
67	TITANIO	LMSNDY LME Comdty	79	YHOO	Yahoo! Inc. (YHOO) -NasdaqGS
68	TM	Toyota Motor Corporation (TM)-NYSE	80	ZEUS	Olympic Steel Inc. (ZEUS) -NasdaqGS
69	TRIGO	W Z2 COMB Comdty	81	ZINC	LMZSDY LME Comdty
70	UPIP	Unwired Planet, Inc. (UPIP) -NasdaqGS			

Anexo 1.C Diagonal de las matrices de covarianza empírica para los periodos 2000-2008 y 2008-2012.

Activo	Diagonal de varianzas 2000-2008	Diagonal de varianzas 2008-2012
A	0.000449	0.0006110
AA	0.000489	0.0006966
AAPL	0.000218	0.0003806
ACEITE DE PALMA	0.000307	0.0005318
ADM	0.000349	0.0005159
AF	0.000400	0.0004342
AIG	0.000126	0.0000958
ALGODÓN	0.000354	0.0002966
ALUMINIO	0.000387	0.0002925
AMGN	0.000288	0.0005761
AMZN	0.000567	0.0007035
ANNB	0.000344	0.0007332
ANSS	0.000243	0.0004560
APC	0.000157	0.0002788
ARMH	0.000439	0.0005337
ARROZ	0.000351	0.0002535
BAC	0.000278	0.0005988
BBY	0.000106	0.0001758
BLK	0.001694	0.0017318
BSX	0.001896	0.0007651
C	0.000432	0.0003420
CACAO	0.000712	0.0017240
CELG	0.001010	0.0021881
CHK	0.002049	0.0010110
CIENA	0.000579	0.0004378
CLF	0.000530	0.0003365
CMC	0.002715	0.0010459
COBRE	0.001425	0.0019264
COST	0.004255	0.0020374
CRMT	0.000259	0.0004116
CRUS	0.000406	0.0010999
CSTR	0.000680	0.0020818
CWTR	0.002611	0.0018537
D	0.000377	0.0008066
DECK	0.001100	0.0005651
DELL	0.000145	0.0005694
DUK	0.001332	0.0004260
EBAY	0.000670	0.0002947
ELNK	0.000537	0.0003921
EMN	0.000771	0.0011480
EXPD	0.001619	0.0012418

Anexo 1.D Diagonal de las matrices de covarianza empírica para los periodos 2000-2008 y 2008-2012.

Activo	Diagonal de varianzas 2000-2008	Diagonal de varianzas 2008-2012
F	0.0010518	0.0006455
FSYS	0.0025950	0.0008780
GANADO	0.0003379	0.0002336
GOMA	0.0027278	0.0022572
IBM	0.0006545	0.0013521
JNPR	0.0010372	0.0005594
JUGO DE NARANJA	0.0003386	0.0009888
L	0.0013401	0.0003625
LSCC	0.0019600	0.0006352
M	0.0006486	0.0010988
MADERA	0.0006845	0.0006501
MSFT	0.0014218	0.0018959
NVE	0.0004205	0.0001923
ORCL	0.0019936	0.0045918
ORO	0.0009802	0.0008822
PALADIO	0.0024746	0.0016878
PFH	0.0008451	0.0010079
PLATA	0.0006757	0.0002679
PLATINO	0.0012500	0.0013845
PLOMO	0.0012448	0.0027411
PNRA	0.0045010	0.0016178
R	0.0007539	0.0014131
RFMD	0.0025057	0.0005256
RIMM	0.0003849	0.0066676
SAN	0.0007831	0.0007462
TITANIO	0.0003943	0.0008584
TM	0.0012378	0.0007859
TRIGO	0.0004973	0.0023536
UPIP	0.0033034	0.0009137
USG	0.0015241	0.0004987
USNA	0.0005572	0.0010049
UTX	0.0012471	0.0005505
VOD	0.0007603	0.0027071
WDC	0.0016716	0.0007871
WLT	0.0005462	0.0003139
X	0.0003711	0.0105312
XOM	0.0005850	0.0010690
YHOO	0.0003575	0.0005998
ZEUS	0.0007177	0.0012043
ZINC	0.0011463	0.0006302

Anexo 1.E Diagonal de la matriz de covarianzas empírica para el periodo 2000-2012.

Activo	Diagonal de varianzas 2000-2012
A	0.000510
AA	0.000567
AAPL	0.000280
ACEITE DE PALMA	0.000392
ADM	0.000412
AF	0.000413
AIG	0.000115
ALGODÓN	0.000332
ALUMINIO	0.000351
AMGN	0.000397
AMZN	0.000619
ANNB	0.000491
ANSS	0.000324
APC	0.000203
ARMH	0.000475
ARROZ	0.000314
BAC	0.000399
BBY	0.000132
BLK	0.001708
BSX	0.001468
C	0.000398
CACAO	0.001095
CELG	0.001456
CHK	0.001656
CIENA	0.000525
CLF	0.000457
CMC	0.002083
COBRE	0.001614
COST	0.003415
CRMT	0.000316
CRUS	0.000669
CSTR	0.001210
CWTR	0.002325
D	0.000539
DECK	0.000897
DELL	0.000306
DUK	0.000989
EBAY	0.000528
ELNK	0.000482
EMN	0.000914
EXPD	0.001476

Anexo 2.A Diagonal de las matrices de covarianza empírica filtradas para los periodos 2000-2008 y 2008-2012.

Activo	Diagonal de varianzas 2000-2008	Diagonal de varianzas 2008-2012
A	0.000445	0.000549
AA	0.000443	0.000623
AAPL	0.000198	0.000317
ACEITE DE PALMA	0.000327	0.000530
ADM	0.000340	0.000504
AF	0.000375	0.000405
AIG	0.000126	0.000092
ALGODÓN	0.000319	0.000268
ALUMINIO	0.000403	0.000292
AMGN	0.000378	0.000502
AMZN	0.000596	0.000605
ANNB	0.000389	0.000626
ANSS	0.000312	0.000414
APC	0.000171	0.000255
ARMH	0.000499	0.000538
ARROZ	0.000382	0.000221
BAC	0.000351	0.000579
BBY	0.000126	0.000160
BLK	0.001492	0.001463
BSX	0.001490	0.000600
C	0.000350	0.000282
CACAO	0.000713	0.001554
CELG	0.001016	0.001860
CHK	0.001581	0.000907
CIENA	0.000468	0.000349
CLF	0.000514	0.000291
CMC	0.002225	0.000879
COBRE	0.001258	0.001776
COST	0.003521	0.001865
CRMT	0.000230	0.000340
CRUS	0.000349	0.000997
CSTR	0.000606	0.001763
CWTR	0.002060	0.001585
D	0.000345	0.000692
DECK	0.000933	0.000488
DELL	0.000109	0.000450
DUK	0.000939	0.000378
EBAY	0.000575	0.000239
ELNK	0.000435	0.000321
EMN	0.000790	0.000909
EXPD	0.001333	0.001158

Anexo 2.B Diagonal de las matrices de covarianzas empírica filtradas para los periodos 2000-2008 y 2008-2012.

Activo	Diagonal de varianzas 2000-2008	Diagonal de varianzas 2008-2012
F	0.001106	0.000487
FSYS	0.001892	0.000817
GANADO	0.000270	0.000201
GOMA	0.002512	0.001906
IBM	0.000581	0.001041
JNPR	0.001006	0.000469
JUGO DE NARANJA	0.000298	0.001002
L	0.001019	0.000319
LSCC	0.001548	0.000564
M	0.000524	0.001424
MADERA	0.000543	0.000582
MSFT	0.001255	0.001822
NVE	0.000496	0.000156
ORCL	0.001843	0.005130
ORO	0.000808	0.000857
PALADIO	0.002046	0.001564
PFH	0.000808	0.000846
PLATA	0.000476	0.000225
PLATINO	0.001338	0.001149
PLOMO	0.001346	0.002152
PNRA	0.004554	0.001499
R	0.000607	0.001183
RFMD	0.001947	0.000446
RIMM	0.000324	0.006736
SAN	0.000696	0.000641
TITANIO	0.000351	0.000689
TM	0.000997	0.000767
TRIGO	0.000463	0.001875
UPIP	0.002287	0.000796
USG	0.001265	0.000422
USNA	0.000545	0.000858
UTX	0.001263	0.000463
VOD	0.000719	0.002252
WDC	0.001340	0.000672
WLT	0.000428	0.000261
X	0.000310	0.008039
XOM	0.000512	0.000891
YHOO	0.000349	0.000484
ZEUS	0.000564	0.000981
ZINC	0.000859	0.000598

Anexo 2.C Diagonal de la matriz de covarianza empírica filtrada para el periodo 2000-2012.

Activo	Diagonal de varianzas 2000-2012
A	0.000519
AA	0.000561
AAPL	0.000282
ACEITE DE PALMA	0.000438
ADM	0.000427
AF	0.000403
AIG	0.000111
ALGODÓN	0.000308
ALUMINIO	0.000346
AMGN	0.000450
AMZN	0.000632
ANNB	0.000539
ANSS	0.000373
APC	0.000219
ARMH	0.000528
ARROZ	0.000328
BAC	0.000481
BBY	0.000153
BLK	0.001596
BSX	0.001155
C	0.000346
CACAO	0.001181
CELG	0.001550
CHK	0.001338
CIENA	0.000453
CLF	0.000421
CMC	0.001674
COBRE	0.001587
COST	0.002846
CRMT	0.000313
CRUS	0.000714
CSTR	0.001300
CWTR	0.001984
D	0.000561
DECK	0.000759
DELL	0.000325
DUK	0.000724
EBAY	0.000439
ELNK	0.000415
EMN	0.000931
EXPD	0.001309

Anexo 2.D Diagonal de la matriz de covarianza empírica filtrada para el periodo 2000-2012.

Activo	Diagonal de varianzas 2000-2012
F	0.000848
FSYS	0.001463
GANADO	0.000255
GOMA	0.002360
IBM	0.000928
JNPR	0.000775
JUGO DE NARANJA	0.000647
L	0.000725
LSCC	0.001134
M	0.000871
MADERA	0.000603
MSFT	0.001583
NVE	0.000331
ORCL	0.003300
ORO	0.000861
PALADIO	0.001901
PFH	0.000888
PLATA	0.000390
PLATINO	0.001317
PLOMO	0.001942
PNRA	0.003073
R	0.000992
RFMD	0.001296
RIMM	0.003487
SAN	0.000716
TITANIO	0.000583
TM	0.000918
TRIGO	0.001336
UPIP	0.001710
USG	0.000904
USNA	0.000753
UTX	0.000895
VOD	0.001679
WDC	0.001088
WLT	0.000378
X	0.005028
XOM	0.000771
YHOO	0.000458
ZEUS	0.000867
ZINC	0.000775

Anexo 2.E Valores propios de las matrices de correlación mediante TMA de los diferentes rendimientos para los periodos 2000-2008 y 2008-2012.

No.	Valores propios para el periodo 2000-2008	Valores propios para el periodo 2008-2012
A	11.6481294	27.9634144
AA	4.0755949	4.7613057
AAPL	2.9143540	1.8144873
ACEITE DE PALMA	1.7952451	1.6496820
ADM	1.6136166	1.5364380
AF	1.5145351	1.3544277
AIG	1.3788967	1.3120686
ALGODÓN	1.2896738	1.2182824
ALUMINIO	1.2700857	1.1735421
AMGN	1.2344936	1.0998573
AMZN	1.1733584	1.0550830
ANNB	1.1719720	1.0403125
ANSS	1.1587900	1.0119199
APC	1.1291038	0.9841636
ARMH	1.1045495	0.9687416
ARROZ	1.0994371	0.9600042
BAC	1.0868806	0.9504880
BBY	1.0660945	0.9181534
BLK	1.0541886	0.9136228
BSX	1.0506880	0.8565016
C	1.0270154	0.8429980
CACAO	1.0210472	0.8236239
CELG	1.0183812	0.8067393
CHK	1.0017181	0.7931958
CIENA	0.9773383	0.7833166
CLF	0.9721372	0.7678441
CMC	0.9540361	0.7330865
COBRE	0.9324667	0.7126274
COST	0.9269762	0.7016410
CRMT	0.9159989	0.6811763
CRUS	0.8978755	0.6709696
CSTR	0.8890309	0.6650256
CWTR	0.8810702	0.6555166
D	0.8775288	0.6337590
DECK	0.8645183	0.6284661
DELL	0.8436416	0.6050740
DUK	0.8318047	0.5826308
EBAY	0.8193262	0.5747943
ELNK	0.8107016	0.5719955
EMN	0.8036518	0.5665988

Anexo 2.F Valores propios de las matrices de correlación mediante TMA de los diferentes rendimientos para los periodos 2000-2008 y 2008-2012.

No.	Valores propios para el periodo 2000-2008	Valores propios para el periodo 2008-2012
EXPD	0.7933532	0.5486394
F	0.7692048	0.5324127
FSYS	0.7649856	0.5239061
GANADO	0.7482178	0.5179264
GOMA	0.7309275	0.5025596
IBM	0.7211590	0.4827672
JNPR	0.7030126	0.4687138
JUGO DE NARANJA	0.6977419	0.4605311
L	0.6913664	0.4483371
LSCC	0.6790849	0.4335349
M	0.6709131	0.4261657
MADERA	0.6616090	0.4079995
MSFT	0.6570054	0.4016692
NVE	0.6400606	0.3937398
ORCL	0.6355193	0.3918540
ORO	0.6261306	0.3674927
PALADIO	0.6101291	0.3568286
PFH	0.6044304	0.3499408
PLATA	0.5965191	0.3334329
PLATINO	0.5872746	0.3291489
PLOMO	0.5767796	0.3224728
PNRA	0.5687301	0.3173717
R	0.5625269	0.3078581
RFMD	0.5598001	0.2932142
RIMM	0.5538095	0.2915791
SAN	0.5356436	0.2841719
TITANIO	0.5117332	0.2723111
TM	0.5026003	0.2603505
TRIGO	0.4939099	0.2589844
UPIP	0.4815663	0.2549587
USG	0.4697474	0.2425290
USNA	0.4646114	0.2334990
UTX	0.4372695	0.2313001
VOD	0.4267340	0.2166990
WDC	0.4115738	0.2069008
WLT	0.3897844	0.1926195
X	0.3616557	0.1839184
XOM	0.3257517	0.1693346
YHOO	0.2644438	0.1657146
ZEUS	0.2332698	0.1570233

Anexo 2.G Valores propios de las matrices de correlación mediante GH de los diferentes rendimientos para los periodos 2000-2008 y 2008-2012.

No.	Valores propios para el periodo 2000-2008	Valores propios para el periodo 2008-2012
A	25.3933617	60.9612988
AA	6.8029181	7.9474956
AAPL	5.4186811	3.3736904
ACEITE DE PALMA	5.0067872	4.6008240
ADM	4.6902482	4.4659156
AF	4.0815063	3.6500344
AIG	3.5826719	3.4090380
ALGODÓN	3.2629564	3.0823316
ALUMINIO	3.1700136	2.9290499
AMGN	3.0852448	2.7487619
AMZN	2.8644356	2.5756984
ANNB	2.9278773	2.5989591
ANSS	2.6341838	2.3003159
APC	2.5125656	2.1900339
ARMH	2.4219027	2.1241220
ARROZ	2.2720692	1.9839207
BAC	2.0736938	1.8134660
BBY	2.0007793	1.7231327
BLK	1.9813093	1.7171210
BSX	1.8998485	1.5487216
C	1.8058134	1.4822534
CACAO	1.8033489	1.4546646
CELG	1.7803631	1.4103647
CHK	1.7501357	1.3858193
CIENA	1.6578384	1.3287235
CLF	1.6449052	1.2992310
CMC	1.5701607	1.2065200
COBRE	1.5379017	1.1753245
COST	1.5406302	1.1661241
CRMT	1.4456944	1.0750807
CRUS	1.3938266	1.0415868
CSTR	1.3550450	1.0136201
CWTR	1.3336898	0.9922658
D	1.3398785	0.9676720
DECK	1.2926671	0.9397111
DELL	1.2706252	0.9113138
DUK	1.2376112	0.8668746
EBAY	1.2109612	0.8495439
ELNK	1.1907463	0.8401383
EMN	1.2004189	0.8463316

Anexo 2.H Valores propios de las matrices de correlación mediante GH de los diferentes rendimientos para los periodos 2000-2008 y 2008-2012.

No.	Valores propios para el periodo 2000-2008	Valores propios para el periodo 2008-2012
EXPD	0.7349533	0.8170431
F	0.7029715	0.7622207
FSYS	0.6704384	0.7255773
GANADO	0.6434219	0.6952005
GOMA	0.6108101	0.6670122
IBM	0.5875641	0.6619709
JNPR	0.5642150	0.6240804
JUGO DE NARANJA	0.5454669	0.6043527
L	0.5065027	0.5774439
LSCC	0.4763029	0.5396678
M	0.4714118	0.5270582
MADERA	0.4465136	0.4980465
MSFT	0.4280203	0.4872958
NVE	0.3974847	0.4631575
ORCL	0.3906894	0.4554654
ORO	0.3751429	0.4252614
PALADIO	0.3592761	0.4066311
PFH	0.3424556	0.3910953
PLATA	0.3248555	0.3649391
PLATINO	0.3016078	0.3558028
PLOMO	0.2956100	0.3486568
PNRA	0.2862927	0.3345155
R	0.2682539	0.3127742
RFMD	0.2354746	0.2729800
RIMM	0.2321168	0.2735947
SAN	0.2194266	0.2689800
TITANIO	0.1984399	0.2436885
TM	0.1824068	0.2294322
TRIGO	0.1778207	0.2296701
UPIP	0.1675904	0.2140526
USG	0.1543688	0.2020308
USNA	0.1368072	0.1772390
UTX	0.1225301	0.1653004
VOD	0.1149877	0.1584775
WDC	0.1047957	0.1493006
WLT	0.0916885	0.1397553
X	0.0651853	0.1248114
XOM	0.0576885	0.1173399
YHOO	0.0390556	0.1248776
ZEUS	0.0358718	0.1166382
ZINC	0.0187750	0.0723292

Anexo 2.I Vector propio principal bajo TMA asociado al valor propio máximo para los periodos 2000-2008 y 2008-2012.

No.	Vector propio para el 2000-2008	Vector propio para el periodo 2008-2012
A	0.0202910	-0.0319073
AA	0.0105544	-0.0648873
AAPL	0.0137709	-0.0258166
ACEITE DE PALMA	0.0348879	0.0875945
ADM	0.0281030	0.0792815
AF	-0.0697697	0.1525610
AIG	0.0214303	0.0800666
ALGODÓN	-0.0006959	-0.1380042
ALUMINIO	-0.2909708	-0.2244214
AMGN	-0.2733811	-0.0771974
AMZN	0.1028197	0.2021924
ANNB	0.2552491	-0.2565786
ANSS	-0.1251012	-0.3232417
APC	-0.1576048	-0.0416683
ARMH	0.0375957	-0.2294837
ARROZ	0.1968283	0.4824845
BAC	0.1266779	0.3703715
BBY	-0.0818864	-0.0926250
BLK	0.0001602	0.2662379
BSX	-0.0106937	-0.0623733
C	-0.0242500	-0.0042946
CACAO	-0.0139579	0.1281272
CELG	-0.2436273	0.0600394
CHK	-0.2371336	-0.0270337
CIENA	0.0161131	-0.1754960
CLF	0.1862038	0.0938113
CMC	0.2407832	0.0793534
COBRE	-0.0835590	0.0587358
COST	0.2666016	0.1006253
CRMT	0.3044970	-0.0325635
CRUS	0.1740410	0.0974255
CSTR	-0.2398678	-0.0481062
CWTR	0.0243214	0.0846406
D	-0.1158689	0.0202926
DECK	0.0004356	-0.0137191
DELL	0.0800433	0.0407284
DUK	0.0895018	-0.0288284
EBAY	-0.1248475	0.0478100
ELNK	0.1793487	0.0133007
EMN	-0.0229817	-0.0121349

Anexo 2.J Vector propio principal bajo TMA asociado al valor propio máximo para los periodos 2000-2008 y 2008-2012.

No.	Vector propio para el 2000-2008	Vector propio para el periodo 2008-2012
EXPD	0.0262472	-0.0938523
F	0.1588498	0.0330052
FSYS	-0.0019856	0.0293555
GANADO	0.0265469	-0.0914884
GOMA	-0.1264022	0.0528491
IBM	0.0551769	0.0062262
JNPR	0.0416466	-0.0385021
JUGO DE NARANJA	0.0503531	-0.0161596
L	-0.0812076	-0.0214861
LSCC	0.0819195	-0.0281678
M	-0.0937290	0.0267145
MADERA	-0.0488582	0.0158983
MSFT	0.0147909	-0.0042094
NVE	-0.0830264	0.0029690
ORCL	0.0373042	-0.0237389
ORO	-0.0163671	-0.0661672
PALADIO	-0.0558631	-0.0196858
PFH	-0.0073015	0.0614610
PLATA	0.0055477	0.0070550
PLATINO	0.0480973	-0.0106179
PLOMO	-0.0256651	0.0083581
PNRA	-0.0043097	0.0170703
R	-0.0536023	0.0158310
RFMD	-0.0207414	0.0643059
RIMM	0.0726672	-0.0230777
SAN	-0.0076994	-0.0357078
TITANIO	-0.0209508	-0.0202474
TM	-0.0090324	0.0064100
TRIGO	0.0666671	0.0239966
UPIP	0.0049971	-0.0156515
USG	-0.0072506	0.0123568
USNA	0.0546598	0.0510089
UTX	-0.0115794	-0.0141267
VOD	0.0122382	0.0232551
WDC	0.0288640	-0.0068718
WLT	0.0053171	-0.0250494
X	0.0371383	0.0425477
XOM	-0.0007583	-0.0331444
YHOO	-0.0112325	0.0081223
ZEUS	0.0036062	0.0019655

Anexo 3.A Parámetros estimados de la distribución multivariada GH para el periodo 2000-2012.

Párametros escalares	α	γ	δ	λ
	7.876635	0.001590	0.030603	-1.715429

Activo	Diagonal de Δ	μ	β
A	1.402306	0.00084685	-0.80232025
AA	1.458210	0.00183660	-1.00772484
AAPL	0.711913	0.00093033	-0.67883398
ACEITE DE PALMA	1.205456	0.00033726	0.00134805
ADM	1.240799	0.00041314	0.06802866
AF	1.150225	0.00016745	0.73018250
AIG	0.339985	0.00068559	-1.47752144
ALGODÓN	0.845312	0.00069814	0.05855623
ALUMINIO	1.179026	0.00043747	0.10710030
AMGN	1.223291	0.00162076	-0.50167596
AMZN	1.779279	0.00122799	0.25926429
ANNB	1.442198	0.00196562	-0.26793281
ANSS	0.889933	0.00183870	-1.55563691
APC	0.604866	0.00088061	0.65127824
ARMH	1.235741	0.00159409	-0.71003402
ARROZ	0.804021	-0.00034089	0.14416320
BAC	1.106812	0.00173367	-0.01603848
BBY	0.362794	0.00114573	-0.88586005
BLK	3.660308	0.00142585	0.09573042
BSX	2.087978	0.00056401	-0.67431618
C	0.528770	0.00089048	-1.38089322
CACAO	2.571588	0.00123248	0.76878765
CELG	2.820180	0.00085035	0.85312615
CHK	2.996822	0.00058433	0.57348440
CIENA	1.013579	0.00009737	0.50592481
CLF	0.630839	0.00042498	0.58705477
CMC	3.596554	0.00070596	0.20841281
COBRE	3.471551	0.00156900	-0.49127295
COST	5.727331	0.00066256	-0.35583773
CRMT	0.692499	0.00038007	0.45773827
CRUS	1.429330	0.00002220	1.40271668
CSTR	2.274145	0.00078088	-0.87990937
CWTR	4.142640	-0.00002009	0.08149142
D	1.102537	-0.00023490	2.64110997
DECK	1.483794	0.00062442	0.78524372
DELL	0.419337	0.00025963	-0.07812492
DUK	1.406887	0.00041761	-0.27578287
EBAY	1.015905	0.00062061	-0.13842482
ELNK	0.728470	0.00013106	0.07666564
EMN	1.516944	0.00084413	-0.67311208
EXPD	3.124928	0.00009714	-0.15830271

Anexo 3.B Parámetros estimados de la distribución multivariada GH para el periodo 2000-2012.

Activo	Diagonal de Δ	μ	β
F	0.657350	0.00133551	-4.72271183
FSYS	3.015065	0.00003151	-0.19090038
GANADO	0.527581	0.00050825	0.24847990
GOMA	5.328495	0.00067106	0.03299351
IBM	1.668161	0.00089193	-1.07672535
JNPR	1.177637	0.00046208	-0.16183405
JUGO DE NARANJA	0.940974	0.00069669	0.10987684
L	1.720229	0.00015748	-0.07536527
LSCC	1.782318	0.00066812	-0.19809950
M	0.580311	0.00023833	1.50602557
MADERA	1.238480	-0.00015765	0.69876250
MSFT	3.155518	0.00252611	-0.59202626
NVE	0.383262	0.00065318	-2.13349927
ORCL	5.527883	-0.00077007	0.35664455
ORO	1.918348	0.00063581	0.37966652
PALADIO	4.287460	0.00028100	0.39697667
PFH	1.884698	0.00057559	0.48376678
PLATA	0.709272	0.00102040	-1.94020632
PLATINO	2.066830	0.00168462	-0.84190926
PLOMO	2.978158	0.00301946	-1.45367537
PNRA	4.601941	-0.00090267	0.82845580
R	1.861126	0.00080039	1.38174621
RFMD	2.041511	0.00175659	-1.38336189
RIMM	1.706690	-0.00054434	2.26660129
SAN	1.329879	0.00022192	-0.36990485
TITANIO	1.094212	0.00029987	1.83674223
TM	1.747843	0.00020170	-0.08593220
TRIGO	1.569670	-0.00006773	-1.11721957
UPIP	2.474591	0.00085734	-0.41430867
USG	1.670531	0.00268816	-1.96943282
USNA	1.419849	0.00145120	-0.56904607
UTX	1.502772	0.00092432	0.17623050
VOD	2.698906	0.00026628	-0.03246846
WDC	2.349506	0.00119359	0.01058598
WLT	0.829566	0.00013359	1.45788783
X	2.364996	-0.00031790	0.34009467
XOM	1.146080	-0.00012666	0.31145922
YHOO	0.920736	0.00050498	0.83820320
ZEUS	1.474150	0.00057763	-0.20614197
ZINC	1.702329	0.00079638	-0.36801466

Anexo 4.A Parámetros estimados de la distribución GH para el periodo 2000-2008.

Parámetros escalares	α	γ	δ	λ
		7.682385	0.001551	0.029848

Activo	Diagonal de Δ	μ	β
A	1.370874	0.00082787	-0.78433642
AA	1.447034	0.00182252	-1.00000169
AAPL	0.699452	0.00091405	-0.66695184
ACEITE DE PALMA	1.199343	0.00033555	0.00134121
ADM	1.237880	0.00041217	0.06786858
AF	1.126331	0.00016398	0.71501444
AIG	0.326327	0.00065805	-1.41816862
ALGODÓN	0.824793	0.00068120	0.05713486
ALUMINIO	1.170396	0.00043427	0.10631639
AMGN	1.196176	0.00158484	-0.49055588
AMZN	1.762901	0.00121669	0.25687781
ANNB	1.408807	0.00192011	-0.26172932
ANSS	0.859476	0.00177577	-1.50239708
APC	0.577592	0.00084090	0.62191121
ARMH	1.210450	0.00156147	-0.69550192
ARROZ	0.798179	-0.00033841	0.14311570
BAC	1.076563	0.00168629	-0.01560015
BBY	0.346170	0.00109322	-0.84526766
BLK	3.615261	0.00140831	0.09455227
BSX	2.055471	0.00055523	-0.66381824
C	0.502848	0.00084682	-1.31319763
CACAO	2.542984	0.00121877	0.76023621
CELG	2.726932	0.00082223	0.82491782
CHK	2.975799	0.00058023	0.56946125
CIENA	1.004277	0.00009647	0.50128212
CLF	0.629737	0.00042423	0.58602953
CMC	3.440243	0.00067528	0.19935495
COBRE	3.450986	0.00155970	-0.48836272
COST	5.557472	0.00064291	-0.34528439
CRMT	0.666304	0.00036569	0.44042314
CRUS	1.429062	0.00002220	1.40245383
CSTR	2.177862	0.00074782	-0.84265571
CWTR	4.036150	-0.00001957	0.07939663
D	1.099835	-0.00023432	2.63463790
DECK	1.483304	0.00062421	0.78498457
DELL	0.414543	0.00025666	-0.07723176
DUK	1.355357	0.00040232	-0.26568167
EBAY	0.982945	0.00060047	-0.13393386
ELNK	0.727916	0.00013096	0.07660733
EMN	1.450273	0.00080703	-0.64352822
EXPD	3.033101	0.00009429	-0.15365092

Anexo 4.B Parámetros estimados de la distribución GH para el periodo 2000-2012.

Activo	Diagonal de Δ	μ	β
F	0.633083	0.00128621	-4.54836726
FSYS	2.966283	0.00003100	-0.18781175
GANADO	0.521972	0.00050285	0.24583812
GOMA	5.249085	0.00066106	0.03250181
IBM	1.611203	0.00086147	-1.03996169
JNPR	1.139202	0.00044700	-0.15655229
JUGO DE NARANJA	0.901500	0.00066747	0.10526748
L	1.694871	0.00015516	-0.07425431
LSCC	1.736719	0.00065103	-0.19303131
M	0.576413	0.00023673	1.49590986
MADERA	1.216859	-0.00015490	0.68656356
MSFT	3.050495	0.00244204	-0.57232223
NVE	0.374176	0.00063769	-2.08292179
ORCL	5.457965	-0.00076033	0.35213366
ORO	1.890537	0.00062659	0.37416232
PALADIO	4.215328	0.00027627	0.39029798
PFH	1.830172	0.00055893	0.46977094
PLATA	0.696984	0.00100272	-1.90659229
PLATINO	2.048306	0.00166952	-0.83436363
PLOMO	2.850497	0.00289003	-1.39136265
PNRA	4.474415	-0.00087765	0.80549809
R	1.813645	0.00077997	1.34649550
RFMD	1.945360	0.00167386	-1.31820825
RIMM	1.648807	-0.00052587	2.18972847
SAN	1.295489	0.00021618	-0.36033944
TITANIO	1.042176	0.00028561	1.74939373
TM	1.672159	0.00019296	-0.08221123
TRIGO	1.491310	-0.00006435	-1.06144646
UPIP	2.417754	0.00083765	-0.40479280
USG	1.662763	0.00267566	-1.96027471
USNA	1.407482	0.00143856	-0.56408968
UTX	1.430744	0.00088001	0.16778382
VOD	2.584872	0.00025503	-0.03109661
WDC	2.301155	0.00116903	0.01036813
WLT	0.816328	0.00013146	1.43462375
X	2.304516	-0.00030977	0.33139734
XOM	1.096196	-0.00012114	0.29790288
YHOO	0.894953	0.00049084	0.81473108
ZEUS	1.443975	0.00056581	-0.20192239
ZINC	1.682468	0.00078709	-0.36372109

Anexo 5.A Parámetros estimados de la distribución GH para el periodo 2008-2012.

Párametros escalares	α	γ	δ	λ
	7.410431	0.001496	0.028792	-1.613896

Activo	Diagonal de Δ	μ	β
A	1.342632	0.00081081	-0.76817796
AA	1.390157	0.00175088	-0.96069545
AAPL	0.683061	0.00089263	-0.65132218
ACEITE DE PALMA	1.183941	0.00033124	0.00132399
ADM	1.222909	0.00040719	0.06704778
AF	1.092493	0.00015905	0.69353325
AIG	0.324836	0.00065504	-1.41168714
ALGODÓN	0.818462	0.00067597	0.05669633
ALUMINIO	1.142296	0.00042384	0.10376388
AMGN	1.165991	0.00154484	-0.47817695
AMZN	1.761782	0.00121592	0.25671476
ANNB	1.342504	0.00182974	-0.24941155
ANSS	0.844297	0.00174441	-1.47586301
APC	0.555102	0.00080816	0.59769634
ARMH	1.155084	0.00149004	-0.66368970
ARROZ	0.788315	-0.00033423	0.14134710
BAC	1.037794	0.00162557	-0.01503836
BBY	0.343295	0.00108415	-0.83824939
BLK	3.571060	0.00139109	0.09339625
BSX	2.045855	0.00055263	-0.66071262
C	0.498665	0.00083978	-1.30227160
CACAO	2.480115	0.00118864	0.74144130
CELG	2.686355	0.00081000	0.81264304
CHK	2.851476	0.00055599	0.54567035
CIENA	0.970127	0.00009319	0.48423616
CLF	0.606191	0.00040837	0.56411786
CMC	3.390080	0.00066543	0.19644807
COBRE	3.405717	0.00153924	-0.48195664
COST	5.466644	0.00063241	-0.33964128
CRMT	0.633164	0.00034750	0.41851817
CRUS	1.408426	0.00002188	1.38220151
CSTR	2.110977	0.00072486	-0.81677662
CWTR	3.971328	-0.00001926	0.07812149
D	1.092439	-0.00023275	2.61692054
DECK	1.433932	0.00060344	0.75885635
DELL	0.396999	0.00024580	-0.07396318
DUK	1.354398	0.00040203	-0.26549379
EBAY	0.937901	0.00057296	-0.12779626
ELNK	0.707906	0.00012736	0.07450138
EMN	1.385183	0.00077081	-0.61464580
EXPD	2.944675	0.00009154	-0.14917142

Anexo 5.B Parámetros estimados de la distribución GH para el periodo 2008-2012.

Activo	Diagonal de Δ	μ	β
F	0.619145	0.00125789	-4.44822946
FSYS	2.873465	0.00003003	-0.18193493
GANADO	0.514119	0.00049528	0.24213961
GOMA	5.056368	0.00063679	0.03130853
IBM	1.554506	0.00083116	-1.00336589
JNPR	1.090342	0.00042783	-0.14983780
JUGO DE NARANJA	0.892135	0.00066053	0.10417393
L	1.623496	0.00014862	-0.07112730
LSCC	1.679863	0.00062972	-0.18671196
M	0.572398	0.00023508	1.48548953
MADERA	1.193170	-0.00015188	0.67319830
MSFT	2.897977	0.00231994	-0.54370743
NVE	0.365464	0.00062285	-2.03442590
ORCL	5.364299	-0.00074728	0.34609055
ORO	1.837796	0.00060911	0.36372416
PALADIO	4.154906	0.00027231	0.38470345
PFH	1.751861	0.00053502	0.44966996
PLATA	0.666848	0.00095937	-1.82415747
PLATINO	1.978408	0.00161255	-0.80589102
PLOMO	2.764242	0.00280258	-1.34926040
PNRA	4.403566	-0.00086376	0.79274365
R	1.728700	0.00074344	1.28343054
RFMD	1.886214	0.00162297	-1.27813015
RIMM	1.615293	-0.00051519	2.14521893
SAN	1.240700	0.00020704	-0.34509998
TITANIO	1.021366	0.00027990	1.71446205
TM	1.660484	0.00019162	-0.08163723
TRIGO	1.442924	-0.00006226	-1.02700747
UPIP	2.401399	0.00083198	-0.40205457
USG	1.606459	0.00258506	-1.89389688
USNA	1.400179	0.00143110	-0.56116291
UTX	1.401453	0.00086200	0.16434876
VOD	2.467668	0.00024347	-0.02968661
WDC	2.196293	0.00111575	0.00989566
WLT	0.799064	0.00012868	1.40428382
X	2.291435	-0.00030801	0.32951636
XOM	1.053254	-0.00011640	0.28623295
YHOO	0.863863	0.00047379	0.78642740
ZEUS	1.414679	0.00055433	-0.19782568
ZINC	1.663976	0.00077844	-0.35972351

Anexo 6.A Pesos de un portafolio de Markowitz bajo diferentes matrices de covarianzas para el periodo 2000-2012.

Activo	Portafolio de Markowitz		
	Ponderaciones usuales	Ponderaciones bajo TMA	Ponderaciones bajo GH
A	0.02638816	0.03082999	0.01545582
AA	0.01197222	0.00965379	0.00044161
AAPL	0.03018794	0.02332954	0.02698626
ACEITE DE PALMA	0.03174762	0.03287188	0.02755423
ADM	0.01505536	0.01198456	0.01506163
AF	0.01617445	0.01440950	0.02850782
AIG	0.17394410	0.20177470	0.11917690
ALGODÓN	0.02975576	0.02875886	0.03022750
ALUMINIO	0.05726845	0.06675320	0.04184051
AMGN	0.00002131	0.00003932	0.00001479
AMZN	0.00001461	0.00002174	0.00002677
ANNB	0.00001101	0.00001757	0.00001210
ANSS	0.00001081	0.00001715	0.00001524
APC	0.01853993	0.02656558	0.00060617
ARMH	0.11340920	0.06424731	0.11619770
ARROZ	0.20400340	0.16687800	0.22847970
BAC	0.00000299	0.00000543	0.00000623
BBY	0.08836798	0.11390100	0.04241328
BLK	0.00000691	0.00000852	0.00000833
BSX	0.00001426	0.00002423	0.00000988
C	0.00005353	0.00004798	0.00003138
CACAO	0.00000367	0.00000511	0.00000572
CELG	0.00000353	0.00000486	0.00000810
CHK	0.00001299	0.00001697	0.00004138
CIENA	0.00014284	0.00005922	0.00067652
CLF	0.00031918	0.00013587	0.00022979
CMC	0.00369650	0.00306121	0.00208633
COBRE	0.00000655	0.00000865	0.00000530
COST	0.00001367	0.00002242	0.00000663
CRMT	0.00068895	0.00018649	0.00308399
CRUS	0.00000719	0.00000815	0.00001476
CSTR	0.00000919	0.00001010	0.00001143
CWTR	0.00000596	0.00000838	0.00000855
D	0.00004324	0.00004980	0.00118822
DECK	0.00645425	0.00571598	0.01747587
DELL	0.04654350	0.05028820	0.10978110
DUK	0.00014599	0.00010808	0.00002101
EBAY	0.00168325	0.00353146	0.00008914
ELNK	0.00046756	0.00094100	0.00017601
EMN	0.00001632	0.00001941	0.00001615
EXPD	0.00000714	0.00001011	0.00000859

Anexo 6.B Pesos de un portafolio de Markowitz bajo diferentes matrices de covarianzas para el periodo 2000-2012.

Activo	Portafolio de Markowitz		
	Ponderaciones usuales	Ponderaciones bajo TMA	Ponderaciones bajo GH
F	0.00001559	0.00001541	0.00001297
FSYS	0.00000892	0.00001258	0.00001119
GANADO	0.02000452	0.02384665	0.01687287
GOMA	0.00000707	0.00000822	0.00001188
IBM	0.00002900	0.00003598	0.00001367
JNPR	0.00013145	0.00007479	0.00006018
JUGO DE NARANJA	0.00001135	0.00001385	0.00002120
L	0.00025626	0.00213494	0.00003895
LSCC	0.00002648	0.00004613	0.00001583
M	0.00735028	0.00722782	0.04850352
MADERA	0.00004117	0.00004311	0.00028036
MSFT	0.00129971	0.00018689	0.00001102
NVE	0.03172795	0.03329358	0.02621980
ORCL	0.00001126	0.00001192	0.00004814
ORO	0.00046650	0.00018269	0.00023198
PALADIO	0.00001184	0.00001108	0.00001437
PFH	0.00040465	0.00012873	0.00521806
PLATA	0.02135626	0.02832487	0.00020162
PLATINO	0.00000463	0.00000674	0.00000513
PLOMO	0.00000320	0.00000419	0.00000324
PNRA	0.00000934	0.00001046	0.00004204
R	0.00000419	0.00000531	0.00001571
RFMD	0.00004464	0.00005191	0.00000981
RIMM	0.00000929	0.00001174	0.00004244
SAN	0.00010127	0.00004277	0.00007855
TITANIO	0.00012923	0.00004345	0.01549342
TM	0.00002932	0.00003219	0.00003248
TRIGO	0.00001427	0.00001715	0.00002775
UPIP	0.00000818	0.00001132	0.00001070
USG	0.00011845	0.00021545	0.00000765
USNA	0.00000607	0.00000659	0.00000860
UTX	0.00077593	0.00021833	0.00006182
VOD	0.01240521	0.01160807	0.02021463
WDC	0.00001796	0.00003440	0.00001469
WLT	0.02578335	0.03559580	0.03767141
X	0.00000396	0.00000509	0.00004065
XOM	0.00007944	0.00009320	0.00007954
YHOO	0.00005123	0.00003594	0.00032532
ZEUS	0.00000356	0.00000519	0.00000656
ZINC	0.00001545	0.00001421	0.00001179

Anexo 7.A Pesos de un portafolio de Markowitz bajo diferentes matrices de covarianzas para el periodo 2000-2008.

Portafolio de Markowitz			
Activo	Ponderaciones usuales	Ponderaciones bajo TMA	Ponderaciones bajo GH
A	0.02638816	0.03082999	0.01545582
AA	0.01197222	0.00965379	0.00044161
AAPL	0.03018794	0.02332954	0.02698626
ACEITE DE PALMA	0.03174762	0.03287188	0.02755423
ADM	0.01505536	0.01198456	0.01506163
AF	0.01617445	0.01440950	0.02850782
AIG	0.17394410	0.20177470	0.11917690
ALGODÓN	0.02975576	0.02875886	0.03022750
ALUMINIO	0.05726845	0.06675320	0.04184051
AMGN	0.00002131	0.00003932	0.00001479
AMZN	0.00001461	0.00002174	0.00002677
ANNB	0.00001101	0.00001757	0.00001210
ANSS	0.00001081	0.00001715	0.00001524
APC	0.01853993	0.02656558	0.00060617
ARMH	0.11340920	0.06424731	0.11619770
ARROZ	0.20400340	0.16687800	0.22847970
BAC	0.00000299	0.00000543	0.00000623
BBY	0.08836798	0.11390100	0.04241328
BLK	0.00000691	0.00000852	0.00000833
BSX	0.00001426	0.00002423	0.00000988
C	0.00005353	0.00004798	0.00003138
CACAO	0.00000367	0.00000511	0.00000572
CELG	0.00000353	0.00000486	0.00000810
CHK	0.00001299	0.00001697	0.00004138
CIENA	0.00014284	0.00005922	0.00067652
CLF	0.00031918	0.00013587	0.00022979
CMC	0.00369650	0.00306121	0.00208633
COBRE	0.00000655	0.00000865	0.00000530
COST	0.00001367	0.00002242	0.00000663
CRMT	0.00068895	0.00018649	0.00308399
CRUS	0.00000719	0.00000815	0.00001476
CSTR	0.00000919	0.00001010	0.00001143
CWTR	0.00000596	0.00000838	0.00000855
D	0.00004324	0.00004980	0.00118822
DECK	0.00645425	0.00571598	0.01747587
DELL	0.04654350	0.05028820	0.10978110
DUK	0.00014599	0.00010808	0.00002101
EBAY	0.00168325	0.00353146	0.00008914
ELNK	0.00046756	0.00094100	0.00017601
EMN	0.00001632	0.00001941	0.00001615
EXPD	0.00000714	0.00001011	0.00000859

Anexo 7.B Pesos de un portafolio de Markowitz bajo diferentes matrices de covarianzas para el periodo 2000-2008.

Activo	Portafolio de Markowitz		
	Ponderaciones usuales	Ponderaciones bajo TMA	Ponderaciones bajo GH
F	0.00001559	0.00001541	0.00001297
FSYS	0.00000892	0.00001258	0.00001119
GANADO	0.02000452	0.02384665	0.01687287
GOMA	0.00000707	0.00000822	0.00001188
IBM	0.00002900	0.00003598	0.00001367
JNPR	0.00013145	0.00007479	0.00006018
JUGO DE NARANJA	0.00001135	0.00001385	0.00002120
L	0.00025626	0.00213494	0.00003895
LSCC	0.00002648	0.00004613	0.00001583
M	0.00735028	0.00722782	0.04850352
MADERA	0.00004117	0.00004311	0.00028036
MSFT	0.00129971	0.00018689	0.00001102
NVE	0.03172795	0.03329358	0.02621980
ORCL	0.00001126	0.00001192	0.00004814
ORO	0.00046650	0.00018269	0.00023198
PALADIO	0.00001184	0.00001108	0.00001437
PFH	0.00040465	0.00012873	0.00521806
PLATA	0.02135626	0.02832487	0.00020162
PLATINO	0.00000463	0.00000674	0.00000513
PLOMO	0.00000320	0.00000419	0.00000324
PNRA	0.00000934	0.00001046	0.00004204
R	0.00000419	0.00000531	0.00001571
RFMD	0.00004464	0.00005191	0.00000981
RIMM	0.00000929	0.00001174	0.00004244
SAN	0.00010127	0.00004277	0.00007855
TITANIO	0.00012923	0.00004345	0.01549342
TM	0.00002932	0.00003219	0.00003248
TRIGO	0.00001427	0.00001715	0.00002775
UPIP	0.00000818	0.00001132	0.00001070
USG	0.00011845	0.00021545	0.00000765
USNA	0.00000607	0.00000659	0.00000860
UTX	0.00077593	0.00021833	0.00006182
VOD	0.01240521	0.01160807	0.02021463
WDC	0.00001796	0.00003440	0.00001469
WLT	0.02578335	0.03559580	0.03767141
X	0.00000396	0.00000509	0.00004065
XOM	0.00007944	0.00009320	0.00007954
YHOO	0.00005123	0.00003594	0.00032532
ZEUS	0.00000356	0.00000519	0.00000656
ZINC	0.00001545	0.00001421	0.00001179