

# **Selección óptima de portafolios basada en cadenas de Markov de primer y segundo orden**

---

**Juan Manuel Gómez R. y José Alfredo Jiménez**

Lecturas de Economía - No. 92. Medellín, enero-junio 2020

Juan Manuel Gómez R. y José Alfredo Jiménez

### **Selección óptima de portafolios basada en cadenas de Markov de primer y segundo orden**

**Resumen:** *En búsqueda de generar estrategias de inversión en pro de maximizar el rendimiento esperado y minimizar el riesgo, se estudian dos modelos de selección de portafolios óptimos. El primero se ajusta usando rendimientos logarítmicos, y en el segundo se emplea análisis de componentes principales (ACP) a estos rendimientos. Luego, para cada uno de ellos se establece su rendimiento ponderado y se crean unas medidas para establecer los estados de las cadenas de Markov de primer y segundo orden. Esto permite pronosticar si los portafolios conformados tendrán comportamientos alcistas o bajistas dadas las probabilidades de los estados de las cadenas de Markov. Se realiza una aplicación usando los retornos de precios de cierre diarios de 21 acciones del COLCAP, para el periodo comprendido desde enero de 2014 a octubre de 2017. Se concluye que en el mercado colombiano un portafolio conformado bajo ACP de los rendimientos tiene una mayor rentabilidad esperada y un menor riesgo a largo plazo, teniendo una precisión de pronóstico del modelo dados los vectores estacionarios de las cadenas de Markov.*

**Palabras clave:** *selección de portafolios; cadena de Markov; análisis de componentes principales; aversión al riesgo; índice bursátil.*

**Clasificación JEL:** *G1, G11, O16.*

### **Optimal portfolio selection based on first and second order Markov chains**

**Abstract:** *Searching for create investment strategies in pursuit of maximizing the expected return on investment and minimizing the risk two models of selection of optimal portfolios are studied. The first portfolio composition model is adjusted using logarithmic returns, and the other uses principal component analysis (PCA) at these returns. Then, for each of them its weighted performance is established and measures are created to establish the states of the first and second order Markov chains, this allows to predict whether the shaped portfolios will have bullish or bearish behaviors given the probabilities of the states of the Markov chains. An application is made using the daily closing price returns of 21 COLCAP shares for the period from January 2014 to October 2017. Concluding that in the Colombian Market a portfolio formed by PCA of the returns has a higher expected profitability and less risk in the long term, having an accuracy of model's forecast according with the stationary vectors of the Markov chains.*

**Keywords:** *portfolio selection; Markov chain; principal component analysis; risk aversion; stock index.*

**JEL Classification:** *G1, G11, O16.*

### **Sélection de portefeuille optimal basée sur les chaînes Markov de premier et deuxième ordre**

**Résumé:** *Lorsqu'il s'agit de rechercher des stratégies d'investissement qui maximisent le rendement attendu et qui minimisent le risque, deux modèles de sélection de portefeuille optimal sont souvent étudiés. Le premier modèle est ajusté à l'aide des rendements logarithmiques, tandis que le second applique l'analyse en composantes principales (ACP) à ces rendements. Ensuite, pour chaque modèle, on établit son rendement pondéré et des mesures sont créées pour établir les états des chaînes de Markov du premier et du deuxième ordre. Cela permet de savoir si les portefeuilles formés auront des comportements haussiers ou baissiers, compte tenu des probabilités des états des chaînes de Markov. Une application est faite en utilisant les rendements quotidiens de clôture de 21 actions COLCAP, pour la période de janvier 2014 à octobre 2017. On conclut que sur le marché colombien un portefeuille formé en ACP a une rentabilité attendue plus élevée et un risque inférieur à long terme. La précision de cette prévision est donnée par les vecteurs stationnaires des chaînes de Markov.*

**Mots clés:** *sélection de portefeuille; chaîne de Markov; analyse des composantes principales; aversion au risque; indice boursier.*

**Classification JEL:** *G1, G11, O16.*

# Selección óptima de portafolios basada en cadenas de Markov de primer y segundo orden

Juan Manuel Gómez R. <sup>a</sup> y José Alfredo Jiménez M. <sup>b</sup>

**–Introducción. –I. Selección óptima de portafolios. –II. Análisis de componentes principales. –III. Metodología para el pronóstico de rendimientos y la selección de portafolios. –IV. Resultados: el comportamiento del mercado colombiano. –V. Cadenas de Markov. –Conclusiones. –Agradecimientos. –Referencias.**

doi: 10.17533/udea.le.n92a02

*Primera versión recibida el 20 de diciembre de 2018; versión final aceptada el 15 de agosto de 2019*

## Introducción

Para analizar los mercados financieros existen dos ramas de estudio, las cuales han evolucionado y mejorado por medio de la investigación científica, el actuar empírico de inversores, los grupos de investigación y las personas afines a la bolsa o a los mercados financieros. Lo anterior, responde a la necesidad de aumentar los rendimientos o utilidades que obtiene el inversionista, al tiempo que reduce el riesgo de perder el capital invertido. Estas dos ramas son el análisis técnico y el análisis fundamental (Martínez et al., 2004), sobre este último se supone que es eficiente a largo plazo y que considera información de la empresa y el comportamiento del mercado. Dado que la empresa es quien cotiza en la bolsa y pone a disposición del mercado las acciones en las cuales se invierte, debe mostrar información de variables macroeconómicas del mercado y variables propias de la empresa como balances, utilidades, estados financieros, regulaciones,

---

<sup>a</sup> Juan Manuel Gómez R.: MSc en Estadística, Universidad Nacional de Colombia, Bogotá D. C., Colombia. Dirección electrónica: jumgomezro@unal.edu.co  
<https://orcid.org/0000-0002-2504-1058>

<sup>b</sup> José Alfredo Jiménez M.: profesor asociado, Departamento de Estadística, Universidad Nacional de Colombia, Bogotá D. C., Colombia. Dirección electrónica: josajimenezm@unal.edu.co  
<https://orcid.org/0000-0002-2391-2809>

su entorno político, etc. Por otra parte, el análisis técnico se basa en el histórico de precios y movimientos de las cotizaciones para poder captar información, realizar predicciones e invertir de una forma adecuada. Las investigaciones que se han hecho de análisis técnico indican que en los precios está implícita toda la información, tanto de las variables de cada título como del mercado; sin embargo Parada (2012) argumenta que estos análisis técnicos son muy subjetivos al observador y que no se consideran factores externos que pueden afectar significativamente los precios a futuro y hacer que estas metodologías no sean robustas, coincidiendo con las hipótesis de McQueen y Throley (1991) años atrás. Por esto, el presente artículo se enmarca dentro del análisis técnico y no del fundamental, ya que la información a analizar serán rendimientos de activos.

Definido el análisis que se va a realizar, surge entonces la pregunta: ¿Cuál es la mejor forma de obtener mayores rendimientos en los mercados bursátiles de una forma segura?<sup>1</sup> Para responder esta pregunta y al cómo aumentar la rentabilidad disminuyendo el riesgo, a través de la historia se han desarrollado una serie de metodologías, unas con fundamento estadístico y otras simplemente con el conocimiento empírico del mercado. Entre dichas teorías están: el modelo de media-varianza de Markowitz (1952, 1959), transformaciones de retornos para incorporar la asimetría y exceso de curtosis que puedan tener las distribuciones, mixturas de distribuciones lognormales en Cecchetti y Sigalotti (2013), modelos bajo la distribución *t* student, modelos gaussianos en Rachev et al. (2002), modelos GARCH, mezclas discretas de distribuciones normales y una gran variedad de modelos desarrollados (véase Hyung & de Vries, 2007).

Sin embargo, problemas como los supuestos de normalidad de los retornos, los efectos de patrones no lineales o los factores externos que impacten el mercado como las burbujas especulativas, las noticias que anuncien aumentos o disminuciones de precios de las acciones, las variaciones temporales o quizás factores ambientales, no sean capaces de captar esta información importante y presenten inconsistencias metodológicas para su adecuada aplicación. Para esto, en este artículo, se considera como una

---

<sup>1</sup> Asociados a factores de riesgo de los portafolios, donde pueden haber inversores con mayor aversión al riesgo que otros.

alternativa el uso de cadenas de Markov de primer o segundo orden acompañado de una metodología de reducción de dimensionalidad (número de activos) usando análisis de componentes principales para los rendimientos diarios de cada una de las acciones que se consideran. Con esto se busca, en primer lugar, evitar el supuesto de normalidad multivariada en el portafolio seleccionado donde las cadenas de Markov propuestas son ergódicas y estacionarias; y, en segundo lugar, poder hacer uso de información de patrones no lineales que contengan las series de retornos (hechos estilizados).

Se considera el histórico de rendimientos de los precios de las acciones del mercado colombiano para ejecutar dos modelos de selección óptima de portafolios (media-varianza de Markowitz) basados en cadenas de Markov, donde se consideran dos medidas principales para el análisis: la rentabilidad y el riesgo. La primera es medida como la utilidad o el retorno de los precios de las acciones en distintos períodos, y la segunda es medida a través de la varianza o desviación estándar de los rendimientos en un período. Considerando estas dos medidas, el objetivo de los inversores es poder maximizar la rentabilidad del portafolio al mismo tiempo que se minimizan las medidas de riesgo; usualmente, si se pretende obtener una alta rentabilidad, se debe buscar invertir en acciones de mayor riesgo o volatilidad, pero, para efectos de lo que se desea realizar en este artículo, se considerará que los inversores son aversos al riesgo y buscan minimizar su riesgo.

Posteriormente, se realiza un análisis de componentes principales, con el fin de reducir la dimensionalidad del número de acciones que se está analizando, dado que el comportamiento de varias de estas acciones son similares, al pertenecer algunas a los mismos sectores productivos y estar inmersas en un mismo mercado, en este caso el colombiano. Por último, con esta reducción de dimensionalidad y con la obtención de las nuevas variables de rendimientos, se ejecutan los modelos de selección óptima de portafolios y se observa qué metodología es mejor al instante de adoptar decisiones de inversión. La decisión se basa en los vectores de probabilidad estacionarios de las cadenas de Markov, los cuales nos indicarán las probabilidades estacionarias de los estados bajistas o alcistas que se definirán más adelante.

## I. Selección óptima de portafolios

### A. Modelo de Markowitz o enfoque de media-varianza

El modelo de media-varianza de Markowitz aparece en el año 1952, dando inicio a la llamada teoría moderna de portafolios. En esta teoría se considera el portafolio o cartera como un todo y, por ello, no se analiza cada acción o título por separado; esto permite obtener una rentabilidad y un riesgo global del portafolio. A su vez, lo anterior facilita la diversificación del riesgo, pues ya no se compromete todo el capital posiblemente en un solo título sino que se busca alcanzar una máxima rentabilidad de todas las acciones compradas en conjunto y, a su vez, minimizar el riesgo de pérdida del capital invertido, el cual se cuantifica con la volatilidad o fluctuación del portafolio conformado. Luego, un inversor averso al riesgo, al tener dos o más portafolios con la misma rentabilidad esperada escogerá el que menos riesgo le implique asumir, es decir, el de menor volatilidad.

En este planteamiento del problema se consideran dos escenarios donde el inversionista preferirá invertir entre dos carteras que tengan: i) La misma rentabilidad –esperada la de menor riesgo–. ii) el mismo nivel de riesgo –la que mayor rentabilidad esperada le proporcione–.

Matemáticamente, se consideran  $m$  títulos valores, donde el objetivo principal es hallar las ponderaciones o proporciones de capital,  $w_k$  para  $k = 1, \dots, m$ , a invertir en cada uno de los títulos, teniendo entonces que  $\sum_k w_k = 1$ , siendo  $w_k$  un valor entre  $[0, 1]$ .

El portafolio con mínima varianza global será el portafolio de ponderaciones  $\vec{w} = [w_k]$  que soluciona el problema de optimización:

$$\min_{\vec{w}} \sigma_p^2 = \min_{\vec{w}} \vec{w}^t \Sigma \vec{w} \quad (1a)$$

$$\text{sujeto a } \vec{w}^t \vec{1}_m = \vec{1}_m^t \vec{w} = 1, \quad (1b)$$

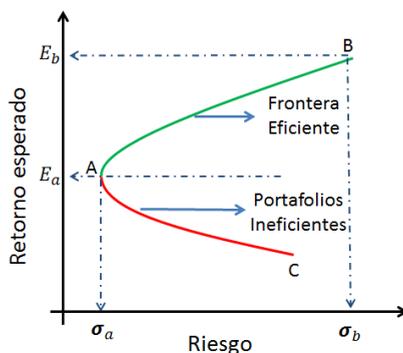
donde  $\Sigma = [\sigma_{ij}]$  es la matriz de varianza-covarianza de los retornos de precios de las acciones. Nótese que el riesgo de la cartera ( $\sigma_p$ ) involucra la varianza de los retornos de cada uno de los títulos que la constituyen y, además, la covarianza de los retornos de cada uno de estos.

## B. Frontera eficiente

Como se mencionó, los modelos de media-varianza consideran dos aspectos importantes: la rentabilidad esperada y el riesgo del portafolio (medido en desviaciones estándares de las rentabilidades de las acciones), estos interactúan entre ellos para poder determinar y diferenciar composiciones de portafolios buenos o malos, o definir qué portafolio es mejor que otro –esto es subjetivo al inversor ya que puede ser averso o no al riesgo–. Tomando la dupla rentabilidad esperada-riesgo y llevándola a un plano cartesiano, se genera la curva denominada frontera eficiente, donde los inversores pueden tomar decisiones y establecer cuál será su rentabilidad media esperada, dependiendo del nivel de riesgo, o, contrariamente, fijar un riesgo para observar qué rentabilidad tendrán.

Suponiendo que los inversores son racionales, en la Figura 1 se observa que para ciertos niveles de riesgo  $\sigma_i$  (desviaciones) existen varias rentabilidades, además el inversor se puede desplazar del portafolio A al portafolio B por la curva AB, sobre la frontera eficiente, buscando así una rentabilidad máxima esperada para cierto nivel de riesgo. Si el inversor quiere obtener la rentabilidad esperada de mínimo riesgo o mínima desviación ( $\sigma_a$ ) obtendrá entonces una rentabilidad de  $E_a$ , pero si él decide asumir un riesgo mayor ( $\sigma_b$ ), obtiene en este caso una rentabilidad esperada mayor ( $E_b$ ).

**Figura 1.** Frontera eficiente



Fuente: elaboración propia.

## II. Análisis de componentes principales

El objetivo de este análisis es obtener la relación que existe entre una agrupación de variables de información observada, buscando sintetizar la información contenida en un conjunto de variables observadas en un menor número de variables no observadas, con la mínima pérdida de información disponible. La finalidad no es buscar relaciones causales o explicativas entre variables, es decir, no se busca predecir o explicar una variable en función de otra(s), sino más bien describir y buscar una adecuada agrupación de variables, véase Peña (2002).

Con este objetivo, lo que se busca es recopilar la información de una agrupación de  $p$  variables observadas, en un conjunto más pequeño de  $r$  nuevas variables, tal que  $r < p$ , donde esté representada la mayoría de la información inicial, generando nuevas variables latentes o no observadas que representan la variabilidad del agrupamiento de datos inicial observado; estas nuevas variables son las componentes principales. Para el desarrollo de este análisis se cuenta con ciertas condiciones y supuestos como:

- Las componentes principales o nuevas variables son combinaciones lineales de las variables iniciales.
- Las componentes principales son incorrelacionadas entre sí, lo que puede facilitar la interpretación de los datos.
- Como se tiene el objetivo de minimizar la pérdida de información, se busca que las nuevas componentes maximicen la correlación global con las variables originales, es decir que, buscando este objetivo, se seleccionan las variables con la máxima variabilidad posible.
- Las variables originales deben tener multicolinealidad, término que se refiere a la existencia de más de una relación lineal y que es usado a menudo en modelos de regresión, en los cuales las variables  $X_1, X_2, \dots, X_p$  pueden estar intercorrelacionados o no de forma perfecta. En algunos textos se recomienda que las correlaciones sean aproximadamente de 0,3 o superior en valor absoluto (véase Jackson, 1991).

En algunos textos se emplea como medida para detectar multicolinealidad, el índice de condición ( $IC$ ), definido como el cociente entre el valor propio más grande y el valor propio más pequeño asociados a la matriz de varianza-covarianza; luego se definen unos rasgos y se detecta el grado de multicolinealidad existente de la siguiente manera (véase Jiménez, 2017).

$$\begin{aligned} 0 \leq IC(\mathbf{X}^t \mathbf{X}) \leq 10 & \text{ no existe multicolinealidad,} \\ 10 < IC(\mathbf{X}^t \mathbf{X}) \leq 30 & \text{ existe multicolinealidad entre moderada y fuerte,} \\ 30 < IC(\mathbf{X}^t \mathbf{X}) & \text{ existe multicolinealidad severa.} \end{aligned}$$

### III. Metodología para el pronóstico de rendimientos y la selección de portafolios

#### *A. Pronósticos bajo cadenas de Markov de primer y segundo orden*

Son distintas las metodologías utilizadas para los análisis de pronóstico de los rendimientos de acciones o títulos valores del mercado bursátil. Muchos de estos métodos, propios de los procesos estocásticos, pueden llevar a soluciones muy complejas y en algunos casos poco convincentes. Dada la ausencia de normalidad de los rendimientos diarios, se tiene como objetivo elaborar modelos robustos que tengan una buena capacidad de pronóstico e integren información dada por sesgos o variabilidad grupal que pueda llegar a generar colas pesadas en las distribuciones empíricas de la muestra, pero que a su vez no refleje la forma leptocúrtica que tenga la distribución. Esto implica que no se tenga un único método que garantice un pronóstico adecuado, pues claramente algunos tienen una mejor capacidad de pronóstico que otros, esto se puede implementar usando medidas que permitan cuantificar esto como el error cuadrático medio.

Los rendimientos pueden ser afectados por patrones no lineales que impacten los precios del mercado bursátil y causen que estos modelos no puedan captar esta información, como puede pasar en los modelos de media-varianza, por ejemplo, utilizados por Markowitz (1952) y demás desarrollos Markowitz (1959), Fabozzi y Markowitz (2002), que usan los dos primeros

momentos de las distribuciones para obtener los pesos de los portafolios, por lo que se ven afectados por estos patrones y, como consecuencia, no obtienen unas estimaciones de los pesos precisas y realistas para los portafolios seleccionados. Dado que pueden existir muchos factores que influyan en los precios de los títulos valores y, por ende, en los rendimientos de estos, es imposible llegar a un modelo único que pronostique y seleccione la mejor cartera considerando las medidas de rentabilidad y riesgo mencionadas previamente; es por esto que se han venido buscando diversas metodologías que sean robustas y tengan aplicabilidad en el entorno de los mercados de valores.

Así, el modelo que se presenta aquí se ha basado en distintas ideas trabajadas desde las últimas tres décadas, con el cual se busca obtener resultados precisos y realistas para tomar decisiones adecuadas y oportunas. Además, se busca capturar los efectos no lineales causados por agentes externos, como pueden ser las decisiones políticas dentro de un país o en el mercado donde se transa el título o valor, así como cambios normativos, factores macroeconómicos que hagan que el mercado en específico tenga cambios que afecten los precios (tales como factores de clima o desastres naturales), especulaciones que logren causar en los inversores incertidumbre o pérdida de confianza al instante de invertir en dichos mercados, etc. Adicionalmente, se pretende evitar los problemas de supuestos relacionados con la no normalidad multivariada en los rendimientos. Desde que Neftci (1984) y Falk (1986) plantearon un enfoque para probar la hipótesis de asimetría en ciclos económicos del mercado norteamericano, se emplean cadenas de Markov de segundo orden definidos en el espacio de estados finitos estacionarios, como sigue:

$$\begin{cases} +1 & \text{si } \Delta X_t > 0 \\ -1 & \text{si } \Delta X_t \leq 0, \end{cases} \quad (2)$$

donde  $\Delta X_t = X_t - X_{t-1}$ . A partir de una matriz de transición  $\mathbf{P} = [p_{ij}]$  con probabilidades de transición:

$$p_{ij} = P(X_{t+1} = j \mid X_t = i) \geq 0; \quad \forall i, j \in S. \quad (3)$$

Estos estudios concluyeron que si  $\mathbf{P}$  era simétrica o los elementos  $p_{ij}$  eran similares, el ciclo económico dado era simétrico, entendiendo este ciclo económico como las fluctuaciones del mercado caracterizadas por períodos de grandes crecimientos y actividad económica seguidos de períodos de baja actividad, crisis o recesiones. Un par de años más tarde McQueen y Throley (1991) quisieron comprobar, desde una metodología similar o extensión de las asimetrías en los ciclos económicos, si los procesos estocásticos relacionados seguían el comportamiento de una caminata aleatoria o si, en realidad, existían patrones en los retornos históricos que definieran una tendencia. Estos autores llegaron a la conclusión de que, en efecto, sí existen tendencias que no son caminatas aleatorias, por lo que rechazan la hipótesis de caminata aleatoria de los retornos a un 1% de nivel de significancia para portafolios equiponderados y a un 5% para portafolios ponderados. Posteriormente Ramírez y Sandoval (2002, 2003) extendieron aún más estos estudios y propusieron un modelo que permitiera pronosticar el comportamiento de los rendimientos de las acciones que transan en la Bolsa Mexicana de Valores (BMV) mediante cadenas de Markov de segundo orden, ante la necesidad de poder pronosticar cuándo las distribuciones empíricas de los rendimientos no son normales estacionarias concluyendo que existía una alta autocorrelación entre las probabilidades de transición del proceso, lo que permite afirmar que existían componentes predecibles no lineales en los rendimientos de la cartera y se generaban así pronósticos fiables.

En concordancia con esto, se definen ciertas medidas a trabajar que recojan información de los rendimientos para distintos rangos de tiempo, los cuales nos darán el fundamento para construir los estados de las cadenas de Markov como se definirá más adelante.

En este artículo, se usan rendimientos logarítmicos, es decir:

$$r_l = \log \left( \frac{P_l}{P_{l-1}} \right), \quad (4)$$

donde  $P_l$  es el precio del valor de la acción o el título valor en el día  $l$ .

Luego se establecen tres estadísticas, las cuales se definen para distintos intervalos de tiempo: diario, semanal, y mensual, que nos servirán para construir los estados y las cadenas de Markov:

$$\overline{R}_l^{(m)} = \frac{1}{m} \sum_{k=1}^m r_{l-k} \quad \text{con} \quad l = 1, 2, \dots \quad \text{y} \quad m \in \{1, 5, 20\}, \quad (5)$$

donde  $m = 1$  indica que la medida es diaria,  $m = 5$  semanal y por último  $m = 20$  mensual, dado que las acciones solo se negocian en los días hábiles de las bolsas de valores.

Dadas las medidas planteadas en (5), se procede a definir los estados de la cadena de primer y segundo orden para cada uno de los intervalos de tiempo definidos anteriormente.

### 1. Estados de primer orden

Se tiene una cadena de Markov con dos estados:  $B \equiv$  la acción baja y  $S \equiv$  la acción sube. El significado de los estados  $S$  y  $B$  dependen de la medida  $\overline{R}_l^{(m)}$  dada en (5) que se haya escogido para la construcción de la cadena, e indicará que, en promedio, los rendimientos de la acción establecidos por (5) fueron positivos ( $S$ ) o negativos ( $B$ ). La matriz de transición es:

$$\begin{array}{c} \text{Estado } i \\ \begin{array}{l} B := \{\overline{R}_{l-1}^{(m)} < 0\} \\ S := \{\overline{R}_{l-1}^{(m)} \geq 0\} \end{array} \end{array} \begin{array}{c} \text{Estado } j \\ B := \{\overline{R}_l^{(m)} < 0\} \quad S := \{\overline{R}_l^{(m)} \geq 0\} \\ \left[ \begin{array}{cc} P(B | B) & P(S | B) \\ P(B | S) & P(S | S) \end{array} \right]. \end{array} \quad (6)$$

### 2. Estados de segundo orden

De acuerdo a Bhat y Miller (2002), las probabilidades de transición en una cadena de Markov de segundo orden vienen dadas por:

$$p_{ijk} = P(X_t = k | X_{t-1} = j, X_{t-2} = i).$$

Usando la misma representación del caso anterior para los estados, se construye la matriz de transición como sigue:

$$\begin{array}{c} \text{Estados } ij \\ \begin{array}{l} BB := \{\bar{R}_{i-2}^{(m)} < 0, \bar{R}_{i-1}^{(m)} < 0\} \\ BS := \{\bar{R}_{i-2}^{(m)} < 0, \bar{R}_{i-1}^{(m)} \geq 0\} \\ SB := \{\bar{R}_{i-2}^{(m)} \geq 0, \bar{R}_{i-1}^{(m)} < 0\} \\ SS := \{\bar{R}_{i-2}^{(m)} \geq 0, \bar{R}_{i-1}^{(m)} \geq 0\} \end{array} \end{array} \left[ \begin{array}{cccc} \text{Estados } jk & & & \\ & BB & BS & SB & SS \\ & \{\bar{R}_{i-1}^{(m)} < 0, \bar{R}_i^{(m)} < 0\} & \{\bar{R}_{i-1}^{(m)} < 0, \bar{R}_i^{(m)} \geq 0\} & \{\bar{R}_{i-1}^{(m)} \geq 0, \bar{R}_i^{(m)} < 0\} & \{\bar{R}_{i-1}^{(m)} \geq 0, \bar{R}_i^{(m)} \geq 0\} \\ & P(B | BB) & P(S | BB) & 0 & 0 \\ & 0 & 0 & P(B | SB) & P(S | SB) \\ & P(B | BS) & P(S | BS) & 0 & 0 \\ & 0 & 0 & P(B | SS) & P(S | SS) \end{array} \right]. \quad (7)$$

Si se tiene un conjunto de  $t$  precios históricos para el análisis, se tiene entonces  $t - 1$  número de rentabilidades diarias,  $\lfloor \frac{t-1}{5} \rfloor$  medidas semanales y  $\lfloor \frac{t-1}{20} \rfloor$  medidas mensuales, donde  $\lfloor a \rfloor$  es la función *parte entera*.

Nótese que estos estados definidos son recurrentes y aperiódicos, ya que la probabilidad de que la cadena retorne a cualquier estado de la cadena es 1 en  $t$  número de días, semanas o meses; muy posiblemente, algunos de estos estados tendrán un menor tiempo esperado de retorno que otros.

### ***B. Rendimientos bajo la composición de portafolios óptimos***

Dado que la composición de los portafolios se conforma respecto al riesgo y al rendimiento esperado de las acciones, se realiza un análisis de todos los títulos valores o acciones considerados, para ver qué tan volátiles o riesgosos son y, paralelamente, ver su rendimiento medio; finalmente se busca encontrar las participaciones óptimas de cada título dentro del portafolio, que permitan obtener: el mayor rendimiento dado un nivel de riesgo o el menor riesgo dado un rendimiento esperado. Puede que bajo las metodologías propuestas de selección del portafolio óptimo algunas acciones sean ponderadas con 0 pesos, los cuales no serán incluidos en la composición del portafolio; también se podría pensar en trabajar con portafolios equiponderados, donde el peso de cada acción sea  $w_j = 1/m$  (donde  $m$  es el total de acciones distintas a incluir en el portafolio), pero esto es poco común e intuitivamente poco lógico para un inversionista, ya que se está distribuyendo una participación igual a un título que puede que tenga un alto riesgo o alta volatilidad y rendimiento medio negativo o tendencia bajista, con un título que tenga un rendimiento medio positivo y sea poco riesgoso.

### 1. *Composición bajo el modelo de Markowitz*

El portafolio óptimo seleccionado se obtiene a partir de la razón *retorno/riesgo* de todos los portafolios de la frontera eficiente y se escoge la razón más grande, la cual será tangente a la frontera eficiente; así se logra maximizar el rendimiento y minimizar la varianza global. Con esto, se configuran los pesos  $w_j$  para cada una de las  $m$  acciones, para definir los rendimientos del nuevo portafolio óptimo como la ponderación de los rendimientos individuales por acción, así:

$$\widehat{r}_i = w_1 r_{i1} + w_2 r_{i2} + \cdots + w_j r_{ij} + \cdots + w_m r_{im} = \sum_{j=1}^m w_j r_{ij}, \quad (8)$$

siendo  $r_{ij}$  el rendimiento en el instante  $i$  de la acción  $j$ , y  $w_j$  el peso en el portafolio óptimo de la acción  $j$ , obtenido por el modelo de Markowitz.

Una vez definidos estos nuevos rendimientos  $\widehat{r}_i$  dados en (8), se determinan las medidas diarias, semanales y mensuales usando (5), para la respectiva configuración de los estados y la estimación de las probabilidades de transición del modelo de Markov.

### 2. *Composición bajo el análisis de componentes principales*

De manera similar a como se configura la selección del portafolio con sus respectivos pesos bajo la composición del modelo de Markowitz, visto en la sección anterior, se realizará un análisis de componentes principales para conformar nuevas variables que permitan resumir la información comprendida en el total de acciones trabajadas; luego, con estas componentes, que se suponen incorrelacionadas, se vuelve a configurar un portafolio óptimo bajo el modelo de Markowitz, para así obtener otra selección del portafolio.

Los primeros desarrollos hechos por Partovi y Caputo (2004), donde se introdujeron las primeras definiciones de *portafolios principales* para analizar el problema de portafolios eficientes y poder disminuir su complejidad, se basaron en la idea de trabajar con conjuntos o acciones que no tuvieran correlación entre sí; para esto trabajaron bajo el análisis de componentes

principales. Estos autores encontraron que en algunos casos que las ponderaciones o pesos para ciertas acciones podían resultar negativas, lo que interpretaron como posibles operaciones en corto. Esta investigación permitió otros distintos desarrollos como los obtenidos por Meucci (2010), quien se propuso evaluar la diversificación de los portafolios en cualquier mercado, no solo analizando la diversificación de un portafolio individual si no la distribución de la diversificación, usando portafolios no correlacionados donde se hizo uso de componentes principales.

Para el desarrollo de este artículo se construyen los portafolios principales con base en Yang (2015), quien conforma portafolios principales para el mercado de acciones australiano, y Pasini (2017), quien se basa en la misma metodología para los títulos del índice industrial americano (Dow Jones Industrial –DJI–), siguiendo una serie de pasos para la construcción de los portafolios principales, los cuales se enuncian a continuación:

- i) Realizar un análisis de componentes principales de los retornos de las acciones, y obtener los coeficientes de estos componentes.
- ii) Al obtener estos coeficientes, se nota que algunos son positivos y otros negativos, luego, cuando las ponderaciones son negativas, se interpreta como que se pueden realizar operaciones en corto.
- iii) Teniendo que  $\mathbf{A}$  es la matriz de coeficientes de los componentes principales, la cual es la misma matriz formada por los vectores propios de la matriz de varianza covarianza, se toma ésta como los pesos para construir los nuevos rendimientos o portafolios principales  $\widehat{\mathbf{R}}_p = \mathbf{A}\mathbf{R}^T$ , donde  $\mathbf{R}^T$  son los rendimientos de las acciones originales y  $\mathbf{A}$  la matriz de los coeficientes de los componentes. Como por construcción estos componentes son ortogonales, se tiene que  $\mathbf{A}^T = \mathbf{A}^{-1}$ , los cuales actúan como los *pesos* de las nuevas variables (véase Jiménez, 2017).
- iv) Una vez teniendo los portafolios principales, se procede a conformar el portafolio óptimo bajo el modelo de Markowitz. Con estos nuevos rendimientos  $\widehat{\mathbf{R}}_p$  como en la expresión (8), y partiendo de estos, como se mencionó anteriormente, se determinan las medidas que permiten construir los estados y las cadenas de Markov de los nuevos rendimientos.

#### **IV. Resultados: el comportamiento mercado colombiano**

Para realizar el análisis, se tomó como referencia el mercado colombiano, representado por el índice bursátil COLCAP, el cual refleja las variaciones de los precios de las acciones más líquidas de la Bolsa de Valores de Colombia (BVC). Analizando el período comprendido entre enero de 2014 y octubre de 2017 (46 meses), se observó que 21 de las 25 acciones que ponderan este índice bursátil en el último trimestre del año 2016 tienen información completa en este rango de tiempo, razón por la cual se trabajó con estos títulos (véase: Tabla 1). Estos datos fueron tomados de la página web de la BVC y se emplearon para el análisis los precios de cierre de cada día hábil durante este rango de tiempo.

Para observar el comportamiento de los precios de cada una de las acciones consideradas, se realizaron tres figuras de los precios al cierre de cada día (véanse Figuras: 2, 3 y 4). Se evidencia que, generalmente, la mayoría de las 21 acciones analizadas tienen un comportamiento muy similar, lo cual se puede interpretar como que, posiblemente, tienen una correlación alta entre ellas, ya que están dentro del mismo mercado, por lo que, los factores que afectan dicho mercado, también impactan a cada una de estas acciones, claro está, dependiendo al sector productivo al que pertenezcan. Además, se ilustra una tendencia alcista en la mayoría de las acciones en el primer semestre y mitad del segundo semestre del año 2014, seguida de una tendencia decreciente hasta inicios del año 2016, y se vuelve a retomar una tendencia alcista en lo transcurrido del año 2016 y gran parte del 2017. Esto evidencia unas caídas significativas en los últimos meses del análisis en algunas de las acciones estudiadas; claramente algunas con mayor variación que las otras, pero en su mayoría coincidiendo con el mismo patrón de comportamiento histórico.

Observando la Figura 2, se puede evidenciar un comportamiento similar en la mayoría de las 8 acciones analizadas, excepto por Ecopetrol, acción que en la mayoría del tiempo analizado tiene una tendencia bajista y luego estable desde el 2016, debido quizás, a la fuerte caída que tuvo el petróleo desde el año 2014 a 2016, recuperándose un poco en el 2017. Sin embargo, como se mencionó anteriormente, las demás acciones tienen una tendencia alcista en

el primer semestre y mitad del segundo semestre del año 2014, seguido de una tendencia bajista hasta inicios del año 2016, para luego retomar una tendencia alcista en el año 2016 y parte del 2017.

**Tabla 1. Conformación COLCAP**

	Nemotécnico	Participación	Descripción
1	PFBLOM	13,12 %	BANCOLOMBIA S.A.
2	GRUPOSURA	9,32 %	GRUPO INVERSIONES SURAMERICANA
3	ECOPETROL	7,5 %	ECOPETROL S.A.
4	NUTRESA	6,76 %	GRUPO NUTRESA S.A.
5	GRUPOARGOS	6,47 %	GRUPO ARGOS S.A.
6	ISA	6,4 %	INTERCONEXION ELECTRICA S.A. E.S.P.
7	BCOLOMBIA	6,29 %	BANCOLOMBIA S.A.
8	PFAVAL	5,54 %	GRUPO AVAL ACCIONES Y VALORES S.A.
9	CEMARGOS	4,98 %	CEMENTOS ARGOS S.A.
10	PFGROUPSURA**	4,82 %	GRUPO INVERSIONES SURAMERICANA
11	EEB	4,38 %	EMPRESA DE ENERGIA DE BOGOTA S.A. E.S.P.
12	PFGROUPARG**	4,24 %	GRUPO ARGOS S.A.
13	PFDVVNDA	3,33 %	BANCO DAVIVIENDA S.A.
14	EXITO	3,04 %	ALMACENES EXITO S.A.
15	BOGOTA	2,81 %	BANCO DE BOGOTA S.A.
16	CORFICOLCF	2,63 %	CORPORACION FINANCIERA COLOMBIANA S.A.
17	PFCEMARGOS**	2,22 %	CEMENTOS ARGOS S.A.
18	CELSIA	1,9 %	CELSIA S.A. E.S.P.
19	CLH**	1,84 %	CEMEX LATAM HOLDING
20	PFAVH	0,97 %	AVIANCA HOLDINGS S.A.
21	CNEC	0,67 %	CANACOL ENERGY LTD
22	CONCRET	0,33 %	CONSTRUCTORA CONCRETO S.A.
23	ETB	0,23 %	EMPRESA DE TELECOMUNICACIONES DE BOGOTA S.A. E.S.P.
24	BVC	0,23 %	BOLSA DE VALORES DE COLOMBIA S.A.

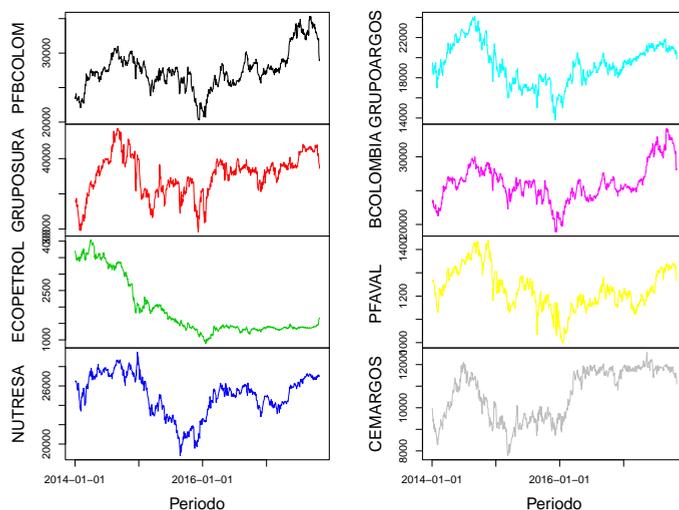
*Nota:* \*\*Valores con información incompleta, excluidos del análisis.

*Fuente:* BVC - tercer trimestre de 2017.

En la Figura 3, se muestra un comportamiento bastante similar al general con una tendencia alcista en el primer semestre y mitad del segundo semestre

del año 2014, seguido de una tendencia bajista hasta inicios del año 2016, para luego retomar una tendencia alcista en el año 2016 y parte del 2017, teniendo unas leves caídas de precios finalizando este último año. Esto, exceptuando a EEB (Empresa de Energía de Bogotá) y CORFICOLCF (Corporación Financiera Colombiana S.A.), donde EEB tienen una tendencia alcista en todo el rango de fechas analizado-quizás por el buen rendimiento de la empresa- y CORFICOLCF evidencia una fuerte caída en el precio de sus acciones en febrero del 2017, según analistas, por una noticia que involucraba a su expresidente en casos de corrupción con Odebrecht.

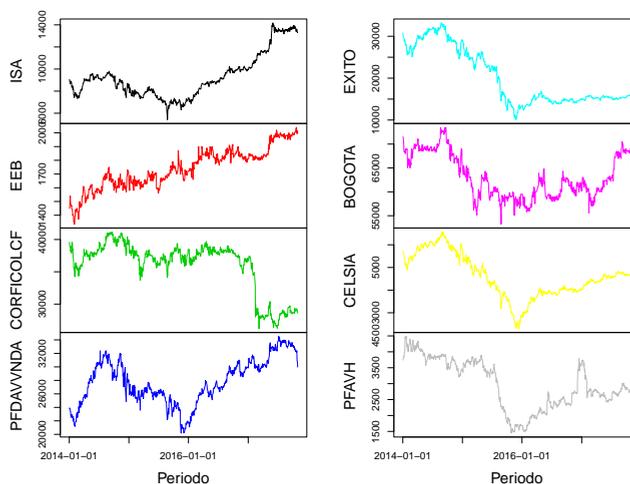
**Figura 2.** Series de precios de cierre diario de acciones que conforman el índice COLCAP



*Fuente:* elaboración propia en R de los precios de cierre diarios.

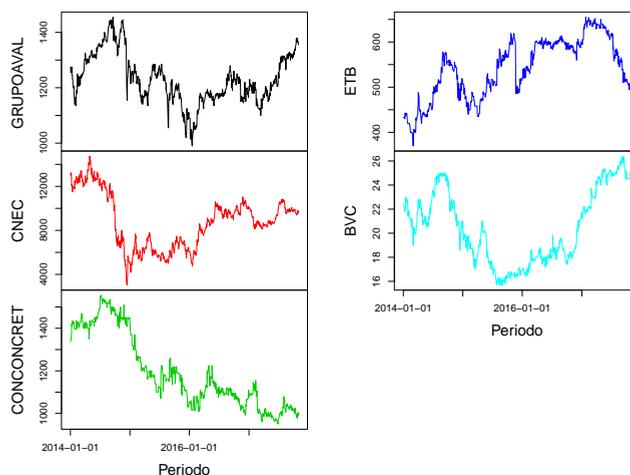
Finalizando el análisis de precios históricos en el rango de tiempo estudiado, se evidencia en la Figura 4, cinco series de precios de acciones, en donde GRUPOAVAL, CNEC y BVC, tienen un comportamiento similar al de las demás acciones, explicado en las Figuras 2 y 3, sin embargo, CONCRET evidencia una tendencia bajista en toda la mayoría del tiempo, estudiado salvo el primer semestre de 2014; caso contrario es el de ETB, el cual generalmente tiene una tendencia alcista, a excepción de ciertos meses donde se revelan unas caídas importantes en el precio de sus acciones.

**Figura 3.** Flujo de relaciones entre las variables Series de precios de cierre diario de acciones que conforman el índice COLCAP



Fuente: elaboración propia en R de los precios de cierre diarios.

**Figura 4.** Series de precios de cierre diario de acciones que conforman el índice COLCAP



Fuente: elaboración propia en R de los precios de cierre diarios.



período observado, y las que presentan un crecimiento anual más alto que las demás.

**Tabla 2.** *Pesos del portafolio*

Nombre	Peso	Nombre	Peso
PFBCOLOM	0,11582	PFDVVNDA	0,02159
GRUPOSURA	0,00000	EXITO	0,00000
ECOPETROL	0,00000	BOGOTA	0,00000
NUTRESA	0,00000	CELSIA	0,00000
GRUPOARGOS	0,00000	PFAVH	0,00000
BCOLOMBIA	0,00000	GRUPOAVAL	0,00000
PFAVAL	0,00000	CNEC	0,00000
CEMARGOS	0,00000	CONCRET	0,00000
ISA	0,33083	ETB	0,07606
EEB	0,45570	BVC	0,00000
CORFICOLCF	0,00000		

*Fuente:* elaboración propia.

### ***B. Portafolio óptimo con aplicación de componentes principales***

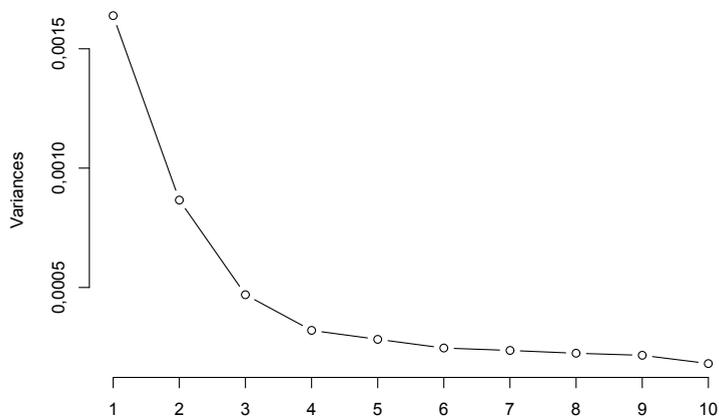
Analizando un poco más en detalle las series, tanto de los precios de las acciones como de los rendimientos, se observan ciertos patrones en las tendencias, que se confirman con las correlaciones tanto en los precios como en los rendimientos. En todos los casos, las correlaciones de los rendimientos son mayores a 0, es decir, tienen un comportamiento directamente proporcional entre las 21 acciones: teniendo una correlación media de 0,247 y algunas de estas llegan a valores de 0,8 de correlación. A partir de este análisis, se realizó una reducción de dimensionalidad del problema, para reducir el número de variables o series de retornos a trabajar, y analizar la estructura y relación que pueden tener estas entre sí. Se tiene

como punto de partida los 916 rendimientos de cada una de las 21 acciones que comprenden el análisis, para así aplicar la metodología estadística de análisis de los componentes principales y construir unas nuevas variables no correlacionadas. Dichas variables, como combinaciones lineales de las originales, explican en gran proporción la información o variabilidad original, posiblemente configurando variables latentes o estructuras de datos que no se puedan ver a simple vista, pero que tengan algún sustento económico. Por ejemplo, podría darse que una componente o nueva variable construida agrupe la información de las acciones de un solo sector económico como el financiero, empresas de servicios, construcción, etc., debido a que tienen un comportamiento parecido en el mercado.

Existen varias reglas para obtener el número de componentes a elegir en el análisis de componentes principales (véase Peña, 2002):

- **Regla 1:** Realizar un gráfico que muestre cada componente con su valor propio asociado o su porcentaje de varianza explicada (gráfico de sedimentación), y gráficamente visualizar un “codo” en la línea trazada, el cual indicará que, desde este punto, los valores propios asociados son aproximadamente iguales y, por ende, su varianza explicada será cada vez menor, tal cual se muestra en la Figura 6.
- **Regla 2:** Determinar desde un principio la varianza explicada que se quiere obtener, por ejemplo el 80 %, 90 % de la varianza explicada total; esta regla puede ser arbitraria y sujeta al investigador.
- **Regla 3:** Seleccionar un número de componentes mayor a una cota, que usualmente se toma como  $\sum \frac{\lambda_i}{p}$ , siendo  $\lambda_i$  cada uno de los valores propios asociados a las componentes y  $p$  el total de componentes.

Para este caso se considera la segunda regla, para explicar como mínimo el 80 % de la varianza. Remitiéndonos a la Figura 6 se trabajará con los primeros 11 componentes que explican un poco más del 80 % de la varianza total. Si, por el contrario, se eligiera la primera o tercera regla, la opción sería considerar solo cuatro componentes que explican únicamente el 55 % de la varianza total (véase Tabla 3).

**Figura 6.** Gráfica de codo ACP - Rendimientos

Fuente: elaboración propia en R.

**Tabla 3.** Variabilidad explicada de las componentes principales

	Desviación estándar	Proporción de varianza	Proporción acumulada
PC1	0,040	0,276	0,276
PC2	0,029	0,146	0,422
PC3	0,022	0,079	0,501
PC4	0,018	0,054	0,555
PC5	0,017	0,048	0,602
PC6	0,016	0,041	0,644
PC7	0,015	0,040	0,683
PC8	0,015	0,038	0,721
PC9	0,015	0,036	0,758
PC10	0,013	0,031	0,788
PC11	0,013	0,028	0,817
PC12	0,013	0,027	0,844

*Continúa*

**Tabla 3.** *Continuación*

	Desviación estándar	Proporción de varianza	Proporción acumulada
PC13	0,012	0,026	0,870
PC14	0,012	0,023	0,892
PC15	0,011	0,021	0,914
PC16	0,011	0,020	0,934
PC17	0,011	0,020	0,953
PC18	0,010	0,015	0,969
PC19	0,009	0,014	0,982
PC20	0,008	0,012	0,994
PC21	0,006	0,006	1,000

*Fuente:* elaboración propia en R.

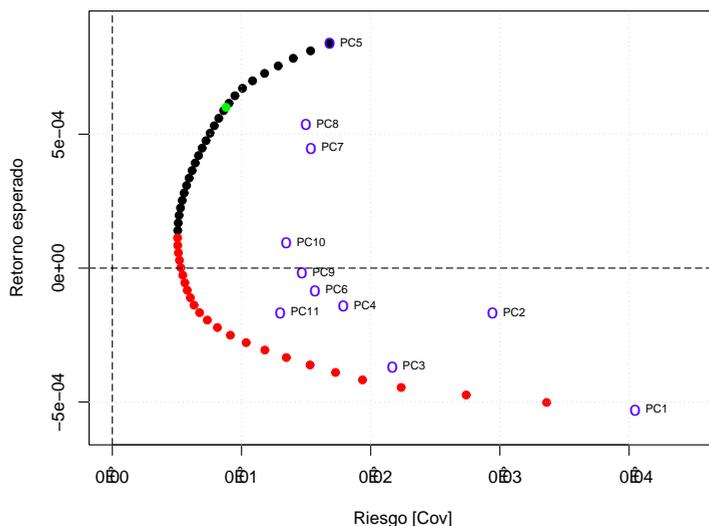
Observando la variabilidad explicada de cada una de las componentes principales configuradas, se evidencia que alrededor del 50 % de la variabilidad de la información es explicada con las primeras 3 componentes, que son aproximadamente el 15 % de las variables originales o dimensionalidad inicial. Un 70 %, 80 % y 90 % de la variabilidad total es explicada con 8, 11 y 15 componentes principales respectivamente, que a su vez corresponden al 38 %, 52 % y 71 % del total de las variables originales.

Teniendo entonces la selección de las 11 primeras componentes principales, se realiza el modelo de Markowitz para obtener el portafolio óptimo tangente a la frontera eficiente, partiendo de las nuevas transformaciones de los retornos dadas como sigue:  $\mathbf{R}_p = \mathbf{E}^{-1}\mathbf{R}^T$ , donde  $\mathbf{E}^{-1}$  es la matriz inversa de los vectores propios, obtenidos al realizar las componentes principales. Estos componentes actúan como *ponderaciones* para el cálculo de los nuevos rendimientos, y adicionalmente esta matriz  $\mathbf{E}^{-1}$  es la misma matriz de rotación de las componentes principales, dado que es una matriz ortogonal. Por otro lado,  $\mathbf{R}^T$  es la matriz transpuesta de los retornos originales. Obteniendo entonces un nuevo portafolio óptimo, con una nueva rentabilidad media estimada y un nuevo riesgo, se calculan las nuevas medidas definidas en la sección III., para la respectiva contrucción de estados y estimación de

las probabilidades de las cadenas de Markov, y se finaliza con un análisis comparativo de las metodologías utilizadas.

En la frontera eficiente (Figura 7) del modelo de Markowitz de los retornos asociados a las componentes principales, se observa que las componentes 5, 7, 8 y 10 se ubican en la figura con rendimientos esperados positivos mientras todos los demás componentes se ubican con rendimientos esperados negativos, lo que da a pensar que el portafolio óptimo, bajo esta metodología, debería estar conformado en mayor proporción por los retornos asociados a las componentes con rendimientos positivos. También se puede observar que los retornos asociados a los componentes PC1, PC2 y PC3 son los que tiene mayor riesgo, al estar a la derecha de la figura, donde en el eje horizontal se tiene la medida de riesgo, lo cual es intuitivamente lógico, ya que son los que más recogen información o variabilidad de los rendimientos originales; además, tienen un rendimiento esperado negativo.

**Figura 7.** Frontera eficiente de los 11 primeros componentes principales



*Fuente:* elaboración propia.

Corroborando las hipótesis anteriormente mencionadas, en la Tabla 4 se observa la composición del portafolio óptimo tangente a la frontera

eficiente, construida por los portafolios que se configuraron bajo el análisis de componentes principales (ACP), donde se pondera en mayor proporción a las componentes 5, 7 y 8. Estas componentes son las de mayor rendimiento y en menor proporción a la componente 10, todas las demás tienen un peso de cero, por lo que quedan por fuera de la composición del portafolio.

**Tabla 4.** Pesos y Riesgo del Portafolio tangente a la frontera eficiente

	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11
Pesos del portafolio	0,0000	0,0000	0,0000	0,0000	0,3824	0,0000	0,2442	0,3072	0,0000	0,0662	0,0000
Riesgo (Covarianza)	0,0000	0,0000	0,0000	0,0000	0,5343	0,0000	0,1819	0,2734	0,0000	0,0103	0,0000

Fuente: elaboración propia en R.

Este nuevo portafolio óptimo tiene un rendimiento esperado de 0,0006, mucho mayor que el rendimiento esperado alcanzado con el portafolio óptimo del modelo de Markowitz con los rendimientos originales, y a su vez tiene una desviación de 0,0938 menor a la anterior como medida de riesgo.

**Tabla 5.** Retorno Esperado y Medidas de Riesgo del Portafolio tangente a la frontera eficiente

Media	Cov	CVar	Var
0,0006	0,0088	0,0183	0,0127

Fuente: elaboración propia en R.

Si se observan los pesos, en la Tabla 6, de cada una de las acciones en las componentes principales configuradas, se nota que:

- Las acciones con mayor ponderación en las componentes 5, 7 y 8, son PFBCOLOM, BCOLOMBIA, PFAVAL, ETB, CEMARGOS, ISA, EEB, BVC, PFDAVNDA, las cuales, de acuerdo a la Figura 5, son las acciones que más rentabilidad positiva tienen.
- Las acciones que tienen una ponderación negativa en estas mismas componentes 5, 7 y 8, son ECOPETROL, ÉXITO, CONCRET, BANCO DE BOGOTÁ, CELSIA, y CORFICOLCF, las cuales, como igualmente se evidencia en la Figura 5 que son las acciones con una

menor rentabilidad esperada (negativa), intuitivamente se interpreta esto como posibles operaciones en corto.

- Si se verifica en la Tabla 6 los valores para las dos primeras componentes, que no tienen ninguna participación en el portafolio óptimo, se evidencian unos pesos mayores para las acciones CNEC, ECOPETROL y ÉXITO, las cuales son las que más volatilidad o riesgo presentan.
- Se tienen unos resultados bastante coherentes con respecto al primer modelo de Markowitz presentado, en cuanto a la escogencia de las acciones para la conformación de los portafolios óptimos, solo que supone entonces que, bajo este modelo y utilizando ACP, se pueden hacer ventas en corto al instante de operar con las acciones.

**Tabla 6.** Pesos de las acciones COLCAP dentro de las componentes principales

Acciones	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	PC11
PFBLOM	0,177	-0,166	0,006	-0,538	0,166	-0,132	0,160	-0,004	0,000	-0,139	-0,077
GRUPOSURA	0,194	-0,182	-0,101	0,021	-0,021	-0,071	0,074	0,076	-0,199	0,022	0,039
ECOPETROL	0,342	-0,022	0,014	-0,122	-0,796	0,374	0,204	-0,051	0,140	-0,110	-0,049
NUTRESA	0,137	-0,097	-0,051	0,024	0,015	-0,001	-0,109	0,005	0,075	0,066	0,089
GRUPOARGOS	0,222	-0,227	-0,145	0,189	0,012	0,014	0,099	0,161	-0,358	0,031	-0,075
BCOLOMBIA	0,187	-0,169	0,016	-0,564	0,194	-0,129	0,163	-0,034	0,017	-0,103	-0,114
PFAVAL	0,106	-0,149	0,003	-0,189	0,154	0,305	-0,091	-0,086	0,209	0,590	0,118
CEMARGOS	0,198	-0,203	-0,136	0,219	0,061	-0,019	0,159	0,208	-0,384	0,026	0,020
ISA	0,183	-0,196	-0,126	0,201	-0,048	-0,107	0,113	-0,065	0,006	0,375	-0,347
EEB	0,073	-0,046	-0,031	0,066	0,045	0,082	-0,010	0,275	0,184	-0,202	0,370
CORFICOLCF	0,124	-0,119	-0,035	0,068	0,018	0,014	-0,062	0,215	-0,146	-0,118	0,075
PFDVAVNDA	0,186	-0,189	-0,066	0,019	0,106	-0,032	-0,030	0,087	0,101	-0,422	0,188
EXITO	0,197	-0,258	-0,119	0,189	-0,125	-0,484	-0,190	-0,642	0,162	-0,093	0,098
BOGOTA	0,098	-0,086	-0,045	0,006	-0,020	0,138	-0,312	0,049	0,141	-0,112	0,276
CELSIA	0,176	-0,154	-0,080	0,126	0,132	0,051	-0,215	0,169	0,292	0,147	0,187
PFAVH	0,209	-0,174	0,941	0,177	0,064	-0,032	0,009	0,009	-0,016	0,006	-0,025
GRUPOAVAL	0,155	-0,140	-0,053	-0,117	0,064	0,105	-0,012	0,045	0,041	0,270	0,132
CNEC	0,639	0,734	-0,026	0,032	0,167	-0,109	-0,053	-0,029	-0,043	0,050	0,018
CONCRET	0,041	-0,049	0,029	-0,187	-0,004	0,338	-0,637	-0,252	-0,517	-0,096	-0,095
ETB	0,057	-0,037	-0,077	0,253	0,422	0,557	0,371	-0,446	0,081	-0,248	-0,063
BVC	0,104	-0,075	-0,076	0,103	0,096	0,065	-0,320	0,267	0,377	-0,201	-0,707

*Fuente:* elaboración propia en R.

## V. Cadenas de Markov

Una vez se establecen los portafolios óptimos bajo los modelos de Markowitz con los rendimientos originales y los rendimientos mediante la metodología de reducción de dimensiones ACP, se consideran los intervalos de tiempo diarios, semanales y mensuales como se definieron en la sección III. A, para poder construir las matrices de transición (véanse las matrices 6 y 7) y establecer los respectivos vectores estacionarios para las cadenas de primer y segundo orden. Así, se obtienen 12 matrices de transición, 6 para cada orden a analizar.

Dadas estas matrices de transición, tanto de primer como de segundo orden, se emplean los teoremas que permiten obtener los vectores de probabilidad estacionarios desde las matrices de transición regulares (véanse Kemeny & Snell, 1976; Jiménez, 2017), tal y como se muestran en las Tablas 7 y 8, donde  $B$  son estados bajistas y  $S$  estados alcistas, como se definió anteriormente en la sección III.

**Tabla 7.** *Vectores estacionarios de primer orden*

Tipo de cadena	B	S
Diario 1er orden	0,465	0,535
Diario 1er orden ACP	0,471	0,529
Semanal 1er orden	0,425	0,575
Semanal 1er orden ACP	0,414	0,586
Mensual 1er orden	0,350	0,650
Mensual 1er orden ACP	0,256	0,744

*Fuente:* elaboración propia.

Como se evidencia en estos 6 vectores estacionarios asociados a cada uno de los modelos de Markov, desarrollados tanto en el primer como en el segundo orden, las entradas de cada uno de estos vectores  $\vec{v}$  se pueden leer o interpretar como la probabilidad estacionaria de estar en dicho estado; por ejemplo, para el modelo con un intervalo de tiempo diario de primer

orden sin usar componentes principales, la probabilidad  $P(S) = 0,535$  es la probabilidad estacionaria de que de un día para otro el precio suba, es decir, que el retorno esperado diario del portafolio seleccionado sea positivo. Así, análogamente, para el modelo con un intervalo de tiempo diario de segundo orden bajo el análisis de componentes principales, la probabilidad  $P(SS) = 0,261$  es la probabilidad estacionaria de que los precios de dos días seguidos suban, o lo que es igual, que dos retornos esperados diarios del portafolio seleccionado sean positivos.

**Tabla 8.** *Vectores estacionarios de segundo orden*

Tipo de cadena	BB	BS	SB	SS
Diario 2do orden	0,226	0,240	0,240	0,294
Diario 2do orden ACP	0,205	0,267	0,267	0,261
Semanal 2do orden	0,182	0,243	0,243	0,332
Semanal 2do orden ACP	0,171	0,243	0,243	0,343
Mensual 2do orden	0,124	0,247	0,247	0,382
Mensual 2do orden ACP	0,076	0,203	0,203	0,518

*Fuente:* elaboración propia.

De igual modo, se puede evidenciar en las Tablas 7 y 8 que si el intervalo de tiempo para calcular los estados de las cadenas, es más amplio, es decir,  $m = 1$ ,  $m = 5$  o  $m = 20$ , diario, semanal o mensual, respectivamente, las probabilidades estacionarias de estar en los estados  $S$  o  $SS$  se vuelven más grandes. Esto se interpreta como que en el mercado colombiano una inversión a largo plazo con estas acciones comprendidas en los portafolios óptimos puede llegar a ser más confiable y rentable que una a corto plazo. Algo semejante ocurre con las cadenas de Markov modeladas para los portafolios resultantes del análisis de componentes principales, el cual arroja probabilidades estacionarias para los estados alcistas  $S$  y  $SS$  mayores que las probabilidades estacionarias de los portafolios que no fueron conformados bajo ACP para las cadenas de primer y segundo orden, excepto para los intervalos de tiempo diarios.

Ahora, se realiza el pronóstico del ajuste del modelo de Markov. Después de realizar la estimación de los vectores estacionarios de las doce matrices de transición para cada cadena de Markov, dependiendo el orden, intervalos de tiempo para la configuración de los estados y estructura de los retornos (originales o transformados por ACP), lo que se quiere ver entonces es el comportamiento de los estados para cada una de las matrices y calcular la proporción de veces que cada estado tiene en los siguientes cuatro meses al análisis, es decir, de noviembre de 2017 a febrero de 2018.

Para esto se conformaron exactamente los mismos portafolios bajo el modelo de Markowitz, tanto de los rendimientos originales, como de los rendimientos transformados bajo el análisis de componentes principales, y partiendo de estos portafolios se calculó la proporción del número de veces que los rendimientos esperados de cada portafolio se encontraron en cada uno de los estados para cada uno de los intervalos de tiempo.

Dicho lo anterior, se observa que los resultados para ver la precisión de pronóstico de los modelos para las cadenas de primer orden (Tabla 9) tienen coherencia y relación con las probabilidades estacionarias de la Tabla 7 estimadas al ejecutar los modelos; igualmente se evidencia esta coherencia para las cadenas de segundo orden en las Tablas 10 contra los vectores de probabilidad estacionarios de la Tabla 8. Adicionalmente, se observa que las proporciones para los estados alcistas como  $S$  para cadenas de primer orden y  $SS$  para cadenas de segundo orden, tienden a ser significativamente más grandes que los otros estados  $B$  (primer orden) o  $BB$ ,  $BS$  y  $SB$  (segundo orden), respectivamente, como se mostró en las estimaciones de los vectores estacionarios; se interpreta entonces que, en general, las acciones seleccionadas de los portafolios óptimos tienen tendencia alcista, y que si en un caso hipotético un inversionista decide contemplar una estrategia de inversión como la planteada, puede tener una utilidad positiva. Hay que mencionar que para los intervalos de tiempo mensuales en las cadenas de segundo orden solo se tienen tres realizaciones de los estados (ya que se toman solo los siguientes 4 meses) donde una es  $BS$ , con una proporción de 0,333 y las otras dos  $SS$ , con una proporción de 0,667.

**Tabla 9.** *Vectores estacionarios. Ajuste de primer orden*

Tipo de cadena	B	S
Diario 1er orden	0,487	0,512
Diario 1er orden ACP	0,418	0,582
Semanal 1er orden	0,467	0,533
Semanal 1er orden ACP	0,400	0,600
Mensual 1er Orden	0,333	0,667
Mensual 1er orden ACP	0,333	0,667

*Fuente:* elaboración propia.

**Tabla 10.** *Vectores estacionarios. Ajuste de segundo orden*

Tipo de cadena	BB	BS	SB	SS
Diario 2do orden	0,287	0,200	0,200	0,312
Diario 2do orden ACP	0,139	0,278	0,291	0,291
Semanal 2do orden	0,267	0,200	0,200	0,333
Semanal 2do orden ACP	0,067	0,333	0,333	0,267
Mensual 2do Orden	0	0,333	0	0,667
Mensual 2do orden ACP	0	0,333	0	0,667

*Fuente:* elaboración propia.

## Conclusiones

Se evidencia, a través de los resultados de los vectores estacionarios asociados a las cadenas de Markov, que los modelos asociados a los retornos, bajo un análisis de componentes principales (ACP), presentan mejores utilidades o rentabilidades esperadas y un riesgo menor a los modelos de Markowitz sin hacer la reducción de dimensionalidad.

Cada vez que se construyen medidas que integren intervalos de tiempo más amplios, como las medidas semanales y/o mensuales, los resultados de

los vectores estacionarios entre estados completamente opuestos ( $B$  vs.  $S$  para cadenas de primer orden,  $BB$  vs.  $SS$  para cadenas de segundo orden) van a tener una diferencia cada vez más grande.

Se concluye que, para el mercado colombiano, una inversión a largo plazo (meses) es más viable, ya que tiene una mayor probabilidad de rentabilidad que una a corto plazo (días), pues las componentes de los vectores estacionarios muestran unas probabilidades mayores para estados alcistas ( $S$ ,  $SS$ ) cuando los intervalos de tiempo son más amplios.

Al observar las proporciones de cada uno de los estados en los 4 meses siguientes al análisis, conocida la realidad de los precios y estados en ese rango de tiempo, bajo las cadenas y temporalidad mencionadas, se concluye que hay coherencia de estas proporciones, con los vectores estacionarios de probabilidad estimados en el modelo, por lo que se observa una precisión en el pronóstico del modelo mediante las cadenas de Markov.

### **Agradecimientos**

Los autores agradecen las observaciones y recomendaciones de los evaluadores anónimos, designados por la revista *Lecturas de Economía*, a la versión preliminar de este artículo, ya que permitieron mejorar y clarificar la metodología planteada en este.

### **Referencias**

- Bhat, U. N. & Miller, G. K. (2002). *Elements of Applied Stochastic Processes* (3ra. ed.). New York: John Wiley & Sons.
- Cecchetti, S. & Sigalotti, L. (2013). *Forward-looking robust portfolio selection* (Temi di Discussione Working Paper, No. 913). Recuperado de: <http://doi.org/10.2139/ssrn.2281906>
- Fabozzi, F. J. & Markowitz, H. M. (2002). *The theory and practice of investment management* (1ra. ed.) New Jersey: John Wiley.

- Falk, B. (1986). Further evidence on the asymmetric behavior of economic time series over the business cycle. *Journal of Political Economy*, 94(5), 1096-1109. <http://doi.org/10.2307/1833193>
- Hyung, N. & de Vries, C. G. (2007). Portfolio selection with heavy tails. *Journal of Empirical Finance*, 14(3), 383-400. <http://doi.org/10.1016/j.jempfin.2006.06.004>
- Jackson, J. E. (1991). *A user's guide to principal components* (1ra. ed.). John Wiley.
- Jiménez Moscoso, J. A. (2017). *Álgebra matricial con aplicaciones en estadística* (3ra. ed.). Bogotá, Colombia: Universidad Nacional de Colombia.
- Kemeny, J. G. & Snell, J. L. (1976). *Finite Markov Chains*. New York: Springer-Verlag.
- Markowitz, H. (1952). Portfolio selection. *The Journal of Finance*, 7(1), 77-91.
- Markowitz, H. (1959). *Portfolio selection: Efficient diversification of investment*. New York: John Wiley.
- Martínez, C. L.; Restrepo, J. & Velásquez, J. (2004). Selección de portafolios usando simulación y optimización bajo incertidumbre. *Dyna*, 71(141), 37-57.
- McQueen, G. & Throley, S. (1991). Are stock returns predictable? A test using Markov chains. *The Journal of Finance*, 46(1), 239-263. <http://doi.org/10.2307/2328695>
- Meucci, A. (2010). *Managing diversification* (Bloomberg Education and Quantitative Research and Education Paper). Recuperado del sitio web de la biblioteca virtual SSRN: [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=1358533](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1358533)
- Neftci, S. N. (1984). Are economic time series asymmetric over the business cycle? *Journal of Political Economy*, 92(2), 307-328. <http://doi.org/10.2307/1831389>

Gómez R. y Jiménez M.: Selección óptima de portafolios basada en cadenas de Markov...

Parada, J. R. (2012). La teoría financiera y análisis técnico. *Revista Universidad EAFIT*, 35(116), 21-36.

Partovi, H. M. & Caputo, M. R. (2004). Principal portfolios: Recasting the Efficient frontier. *Economics Bulletin*, 7(3), 1-10.

Pasini, G. (2017). Principal component analysis for stock portfolio management. *International Journal of Pure and Applied Mathematics*, 115(1), 153-167. <https://ijpam.eu/contents/2017-115-1/12/12.pdf>

Peña, D. (2002). *Análisis de datos multivariantes* (1ra. ed.). Madrid, España: McGraw-Hill Interamericana de España.

Rachev, S.; Huber, I. & Ortobelli, S. (2002). *Portfolio choice with heavy tailed distributions*. Recuperado de: <https://pdfs.semanticscholar.org/41a9/06b013eeb3c89e28435af8375f7eacb53fc3.pdf>

Ramírez, J. C. & Sandoval, R. (2002). *Patrones no lineales en los rendimientos de las acciones de la BMV: Una prueba basada en cadenas de Markov de segundo orden*. Centro de Investigación y Docencia Económicas, A.C. (CIDE). Recuperado de: <http://aleph.academica.mx/jspui/bitstream/156789/4312/1/DOCT2064958.PDF>

Ramírez, J. C. & Sandoval, R. (2003). Una propuesta para evaluar pronósticos de rendimientos de acciones cuando las distribuciones empíricas no son normales estacionarias. *Estudios Económicos*, 18(2), 237-277.

Yang, L. (2015). *An application of principal component analysis to stock portfolio management* (Tesis de maestría, University of Canterbury, Nueva Zelanda).