

Paridad de riesgo jerárquico: aproximación al método y aplicación para el mercado estadounidense

**Hierarchical Risk Parity: Approach to the method
and application for the American market**

Daniel Aragón Urrego*

* Magíster en Finanzas. Docente-Investigador, Observatorio de Economía y Operaciones Numéricas (ODEON), Universidad Externado de Colombia, Bogotá (Colombia). [daniel.aragon@uexternado.edu.co], [ORCID ID: 0000-0001-7677-1379].

Artículo recibido: 27 de mayo de 2022

Aceptado: 12 de julio de 2022

Para citar este artículo:

Aragón-Urrego, D. (2022). Paridad de riesgo jerárquico: aproximación al método y aplicación para el mercado estadounidense. *Odeon*, 21, 105-124.

DOI: <https://doi.org/10.18601/17941113.n21.06>

Resumen

Este artículo presenta el enfoque de paridad de riesgo jerárquico (HRP, por sus siglas en inglés) propuesto por López de Prado para la construcción de portafolios óptimos de inversión empleando aprendizaje no supervisado –*clustering* jerárquico–, que permitan superar algunas limitaciones propias del modelo de Media-Varianza, en particular aquella relacionada con la necesidad de invertir la matriz de covarianzas en el momento de implementar el algoritmo CLA. Se toma una muestra de siete activos del mercado estadounidense, con los cuales se realiza una aplicación del algoritmo HRP propuesto por López de Prado; se encuentra que bajo este modelo la distribución de los activos en diferentes clústeres genera mejoras en términos del retorno esperado, así como del coeficiente de Sharpe en comparación con los resultados del portafolio de media-varianza.

Palabras clave: portafolio óptimo; paridad de riesgo; *clustering*.

Clasificación JEL: C61, G11.

Abstract

This paper presents the Hierarchical Risk Parity (HRP) approach proposed by López de Prado (2016, 2018, 2020) for the construction of optimal investment portfolios using unsupervised learning, hierarchical clustering, which allow overcome some limitations of the Mean-Variance (MV) model, in particular those related to the need to invert the covariance matrix when implementing the CLA algorithm. A sample of 7 assets from the American market is taken, with which an application of the HRP algorithm proposed by López de Prado is carried out, finding that under this model the distribution of assets in different clusters generates improvements in terms of the expected return, as well as of the Sharpe coefficient compared to the results of the Mean-Variance portfolio.

Key words: Optimal portfolio; risk parity; clustering.

JEL classification: C61, G11.

Introducción

La gestión moderna de portafolios ha evolucionado en las últimas décadas para superar algunas limitaciones propias del modelo tradicional de media-varianza (MV) de Markowitz. En este aspecto, siguiendo a López de Prado (2020), el algoritmo de línea crítica (CLA, por sus siglas en inglés) de Markowitz, mediante el cual se construye la frontera eficiente de portafolios óptimos, en la práctica tiende a generar soluciones concentradas e inestables. El principal problema

del modelo MV es su dependencia sobre la matriz de covarianzas, ya que el ejercicio de optimización del portafolio para generar el vector de pesos depende de la inversa de esta matriz. Por ejemplo, cuando existe una alta relación (o correlación) entre los retornos de los activos, la matriz de covarianzas (o de correlaciones) presenta una alta condición número¹ que hace que los valores de su inversa exploten.

En este sentido, López de Prado (2020) resalta tres aproximaciones que se pueden utilizar para superar esta limitación:

- a) Incorporar información adicional sobre los parámetros estimados de los retornos esperados (media) o las covarianzas, en este caso se destaca el trabajo de Black Litterman (1992), en el cual se redefine el ejercicio de optimización al incorporar criterios subjetivos.
- b) La incorporación de restricciones adicionales, como recomiendan Clarke *et al.* (2002).
- c) Mejorar la estabilidad numérica de la matriz inversa de covarianzas, siguiendo el trabajo de Ledoit y Wolf (2004).

Además, el autor resalta el aporte de López de Prado (2016), quien incorpora la técnica de *machine learning* para la asignación de activos a través del modelo paridad de riesgo jerárquico (HRP, por sus siglas en inglés), mediante el cual no se requiere de la matriz inversa de covarianzas. Asimismo, se encuentra la propuesta de Raffinot (2018), que adicional al modelo HRP, plantea la jerarquización de clústeres y la optimización de portafolios sobre las medidas de riesgo a la baja, por ejemplo, el valor en riesgo condicional (CVaR, por sus siglas en inglés), con el objetivo de lograr mejores rendimientos ajustados al riesgo.

En este trabajo se presenta el modelo HRP para la construcción del portafolio óptimo de inversión, siguiendo los trabajos de López de Prado (2016, 2018, 2020). Para ello, se presentan los elementos teóricos y prácticos para la formulación del problema de optimización y el uso de técnicas computacionales para su implementación, siguiendo este enfoque novedoso. Una característica importante

1 De acuerdo con López de Prado (2016), la condición número de una matriz se define como “el valor absoluto de la relación entre sus valores propios máximos y mínimos (por módulos)” (traducción propia). Así, un mayor número de condición conlleva que mayores errores relativos de los datos de entrada impliquen mayores errores relativos de raíces del sistema lineal.

del HRP es la mejora en el desempeño de los portafolios de inversión, la cual es confirmada a partir de una aplicación sencilla para el mercado estadounidense.

Este trabajo está organizado en tres secciones. En la primera se presenta el modelo para la construcción de un portafolio siguiendo el enfoque HRP. La segunda sección presenta la implementación de la metodología propuesta. Finalmente, se presentan las conclusiones del trabajo.

1. El modelo de paridad de riesgo jerárquico

1.1 Modelo MV y el problema del algoritmo CLA

Markowitz (1952) propuso un modelo para la construcción de portafolios óptimos de inversión, conocido como modelo media-varianza (MV). En este modelo, Markowitz considera N activos riesgosos que son caracterizados teniendo en cuenta la naturaleza aleatoria de sus retornos. Para ello considera el vector aleatorio $R = (R_1, R_2, \dots, R_n)^T$ que representa la serie histórica de retornos de los activos, mientras que $E(R) = \mu$ es el vector $N \times 1$ de retornos esperado (o medias) y V es la matriz $N \times N$ de covarianzas.

Teniendo en cuenta esta caracterización, el problema formulado por Markowitz contempla un problema de optimización desde dos posibles escenarios:

- i) Minimizar la varianza (o la volatilidad) del portafolio sujeto a un nivel dado de retorno.
- ii) Maximizar el retorno esperado del portafolio dado un nivel de riesgo.

Para cualquiera de los escenarios, el autor contempla restricciones de igualdad referidas a la sumatoria de pesos igual a 1 ($\sum_{i=1}^N w_i = 1$), así como de desigualdad, en donde los pesos para cada activo tienen un límite superior e inferior, por ejemplo, $0 \leq w_i \leq 1$, cuando se considera la restricción a los cortos, es decir, solo se permiten pesos no negativos.

De esta manera, si el portafolio de inversión tiene un retorno esperado igual a $E(Rp) = w^T \mu$ y una varianza igual a $\sigma_p^2 = w^T V w$, donde w es el vector $N \times 1$ de pesos, entonces, el problema general de optimización para un inversionista que es averso al riesgo se puede ver como un problema de minimización de la varianza del portafolio:

$$\min_{\{w\}} \frac{1}{2} w' V w \text{ s.a. } \sum_{i=1}^N w_i = 1 \quad (1)$$

En la ecuación (1), la restricción de igualdad $\sum_{i=1}^N w_i = 1$ puede reescribirse de forma matricial como $w^T \mathbf{1} = 1$, donde, $\mathbf{1}$ es un vector $N \times 1$ de unos. La solución del problema de optimización establecido por Markowitz permite obtener el portafolio de mínima varianza que viene dado por:

$$w = \frac{V^{-1} \mathbf{1}}{\mathbf{1}' V^{-1} \mathbf{1}} \quad (2)$$

Como se observa en la ecuación (2), la solución del portafolio de mínima varianza depende exclusivamente de la inversa de la matriz de covarianzas.

Por otra parte, al incorporar la restricción de retorno esperado $E(R_p) = w^T \mu$ en el problema descrito en la ecuación (1) para cada portafolio de mínima varianza, se puede obtener toda la frontera eficiente (FE). Aparte de este aporte de Markowitz, se destaca su contribución para la resolución de ecuaciones cuadráticas en lo que se ha denominado el algoritmo CLA, que puede verse como una formulación analítica alternativa para la construcción de la FE.

Formalmente, el problema de optimización se representa de la siguiente manera:

$$\mathcal{L} = w^T V w + 2\lambda_1 [E(R_p) - w^T \mu] + 2\lambda_2 [1 - w^T \mathbf{1}] \quad (3)$$

En donde se tienen las siguientes condiciones de primer orden para la optimización:

$$\frac{\partial \mathcal{L}}{\partial w} = V w - \lambda_1 \mu - \lambda_2 \mathbf{1} = 0 \quad (4)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_1} = E(R_p) - w^T \mu = 0 \quad (5)$$

$$\frac{\partial \mathcal{L}}{\partial \lambda_2} = 1 - w^T \mathbf{1} = 0 \quad (6)$$

Partiendo de (4) se despeja w y se obtiene:

$$w = \lambda_1 V^{-1} \mu - \lambda_2 V^{-1} 1 \quad (7)$$

En este orden de ideas, (7) representa un sistema de n ecuaciones con $n+2$ incógnitas (w_i , λ_1 y λ_2). Para resolver este sistema se deben combinar las condiciones expresada en (5), (6) y (7), de tal manera que λ_1 y λ_2 se resuelvan primero.

$$\lambda_1 = \frac{ZE(R_p) - Y}{D} \quad (8)$$

$$\lambda_2 = \frac{X - YE(R_p)}{D} \quad (9)$$

Siendo:

$$X = \mu^T V^{-1} \mu$$

$$Y = \mu^T V^{-1} 1$$

$$Z = 1^T V^{-1} 1$$

$$D = XZ - Y^2$$

Se debe recordar que siendo V una matriz definida positiva, entonces existe su inversa (V^{-1}), lo cual implica que X , Y , Z y D sean mayores a cero. Resolviendo el sistema mediante el reemplazo de λ_1 y λ_2 en (7) se obtiene el vector de pesos óptimos que se encuentra determinado por la relación lineal con su retorno esperado $E(R_p)$:

$$w_{opt} = g + hE(R_p) \quad (10)$$

En donde:

$$g = \frac{1}{D} [XV^{-1}1 - YV^{-1}\mu]$$

$$h = \frac{1}{D} [ZV^{-1}\mu - YV^{-1}1]$$

Sin embargo, debe tenerse en cuenta que las soluciones generadas por el CLA presentan inestabilidad en los retornos esperados, como lo señalaron Michaud y Michaud (2007). Los autores demostraron que pequeñas desviaciones en los retornos esperados generan portafolios muy diferentes.

En este sentido, la investigación en este campo se ha reorientado hacia el desarrollo de modelos que estimen la matriz de covarianzas, situación que ha puesto en evidencia su sensibilidad a la inversión de esta matriz, la cual debe ser definida positiva. De esta manera, la matriz de covarianzas está sujeta a grandes errores, cuando la matriz está numéricamente mal condicionada, es decir, cuando tiene una alta condición número (Bayle y López de Prado, 2012). En otras palabras, una alta condición número genera que la matriz de covarianza (o de correlaciones) sea inestable. A esto se suma que la mayor correlación de las inversiones genera una mayor necesidad de diversificación.

Además, se ha señalado que el algoritmo CLA de Markowitz conlleva una serie de advertencias para su uso, donde se destaca la necesidad de invertir la matriz de covarianzas (Vyas, 2019). El cumplimiento de esta condición conlleva que los cambios en las correlaciones del mercado, los activos, generen cambios significativos en el inverso de la matriz, con lo cual la hace inestable a la volatilidad del mercado.

Para superar estos problemas, se adoptan las propuestas de López de Prado (2016, 2018) y Vyas (2019), quienes presentan el enfoque HRP. Dicha implementación se puede llevar a cabo a partir de tres etapas:

- *Tree clustering.*
- Cuasidiagonalización.
- Bisección recursiva.

A continuación, se presentan con detalles cada una de ellas.

1.2 Tree clustering

Como se mencionó, el enfoque HRP reconoce las limitaciones que presenta el CLA, entre ellas el problema que involucra la inversa de la matriz de covarianzas. De acuerdo con López de Prado (2018): “todas las inversas son posibles

sustitutos entre sí ya que las matrices de covarianzas carecen de estructura jerárquica, permitiendo que los pesos tengan un comportamiento variable y no deseado, explicando de esta manera la inestabilidad del CLA” (traducción propia). En este punto, López de Prado plantea que la composición del portafolio se puede analizar a partir de la estructura jerárquica de árbol, donde los pesos se distribuyen de arriba hacia abajo y se reajustan al interior de cada jerarquía.

Siguiendo a López de Prado (2018), la matriz de correlación se define por las siguientes entradas:

$$\rho = \{\rho_{i,j} = 1, \dots, N\} \quad (11)$$

Donde: $\rho_{i,j} = \rho[X_i, X_j]$. Una vez se obtiene la matriz de correlación, se procede a determinar una nueva matriz de distancia (D) que viene dada por:

$$D(X_i, X_j) = \sqrt{0.5(1 - \rho_{\{i,j\}})} \quad (12)$$

Después, se requiere establecer una matriz que contenga los resultados de calcular la distancia euclidiana entre cada par de vectores columna de la matriz D . Esto genera una nueva matriz de distancia aumentada (\bar{D}):

$$\bar{D}(i, j) = \sqrt{\sum_{k=1}^N [D(k, i) - D(k, j)]^2} \quad (13)$$

Donde: $\bar{D}(i, j)$ es una función de la matriz de correlación. Una vez se tiene la matriz \bar{D} , se debe tomar un par de activos (i^*, j^*) que tengan la menor distancia para formar el primer clúster. Esto se muestra en la expresión (14), en donde U corresponde al conjunto de clústeres:

$$U[1] = \arg_{\{i,j\}} \min \bar{D}(i, j) \quad (14)$$

Luego, se requiere una matriz *linkage*, o de enlazamiento, que permita actualizar la matriz de distancia \bar{D} . De esta manera, la distancia entre el primer clúster $U[1]$ y cualquier otro activo se determina por:

$$\bar{D}(i, U[1]) = \min(\bar{D}(i, i^*), \bar{D}(i, j^*)) \quad (15)$$

Este proceso se debe repetir para cada activo del portafolio. De esta manera, cada vez que se forma un nuevo clúster de activos se actualiza la matriz de distancia, y este procedimiento se realiza hasta que quede un solo clúster. Para comprender la implementación de esta metodología se desarrolla el siguiente ejemplo, siguiendo a Vyas (2019).

Si se tienen cinco activos, donde los activos α y β tienen la menor distancia, entonces, se forma un clúster con ellos. La siguiente matriz presenta las distancias (D) obtenidas para todos los activos.

	α	β	γ	δ	ϵ
α	0	15	19	24	30
β	15	0	21	23	33
γ	19	21	0	40	29
δ	24	23	40	0	45
ϵ	30	33	29	45	0

El siguiente paso, conocido como *linkage clustering*, consiste en calcular las distancias de los otros elementos con respecto al clúster recién formado, con lo cual se actualiza la matriz de distancia D . En el enfoque de HRP, se suele emplear el método *simple linkage clustering*, que consiste en relacionar a los vecinos más cercanos.

Ahora bien, se aplica la ecuación (15) para calcular la distancia para los activos γ , δ y ϵ con respecto al clúster (α, β) :

$$\bar{D}(\gamma, U[1]) = \min(\bar{D}(\gamma, \alpha), \bar{D}(\gamma, \beta)) \min(19, 21) = 19$$

$$\bar{D}(\delta, U[1]) = \min(\bar{D}(\delta, \alpha), \bar{D}(\delta, \beta)) \min(24, 23) = 23$$

$$\bar{D}(\epsilon, U[1]) = \min(\bar{D}(\epsilon, \alpha), \bar{D}(\epsilon, \beta)) \min(30, 33) = 30$$

Así, la matriz de distancia D se actualiza como:

	(α, β)	γ	δ	\in
(α, β)	0	19	23	30
γ	19	0	40	29
δ	23	40	0	45
\in	30	29	45	0

En este paso se eliminan las respectivas filas y columnas de los activos α y β y se forma una sola columna y fila para este nuevo clúster (α, β) . Se repite el proceso realizado previamente para los activos δ y \in , lo cual redefine la matriz D :

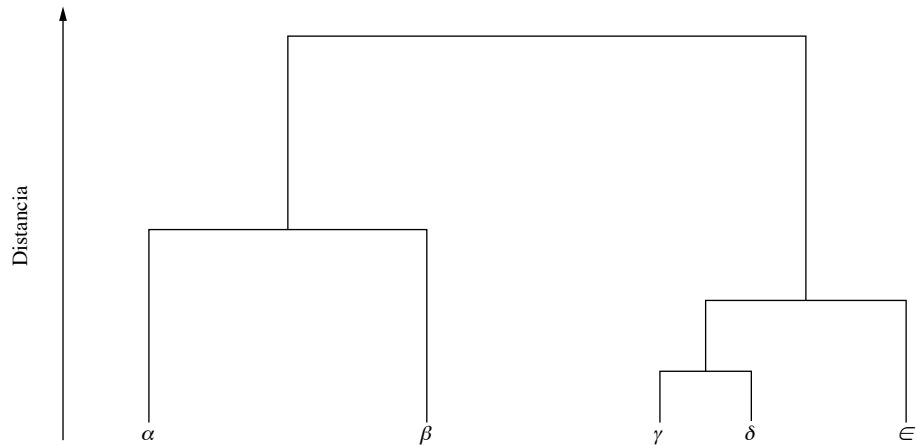
	$(\alpha, \beta), \gamma$	δ	\in
$(\alpha, \beta), \gamma$	0	23	29
δ	23	0	45
\in	29	45	0

Finalmente, al realizar el proceso recursivamente, se obtiene el siguiente resultado:

	$(\alpha, \beta), \gamma, \delta$	\in
$(\alpha, \beta), \gamma, \delta$	0	29
\in	29	0

Gráficamente, el proceso anterior se representa mediante un dendograma, el cual muestra esta estructura de clústeres (figura 1).

Figura 1: Dendograma



Fuente: adaptado de Vyas (2019).

1.3 Cuasi- diagonalización

En esta etapa se realiza el reordenamiento de la matriz de covarianzas de forma semejante al orden que tienen los activos dentro de la matriz *linkage*. De esta forma, a partir de los clústeres formados, se reordenan las columnas y filas de la matriz de covarianzas para que los mayores valores se encuentren cerca y sobre la diagonal, en tanto que las covarianzas más pequeñas se ubiquen fuera de esta.

1.4 Bisección recursiva

En la tercera y última etapa se definen los pesos óptimos para cada activo dentro del portafolio. Para ello, siguiendo a Vyas (2019), se presentan a continuación los respectivos pasos:

- i. Se denotan los pesos de los activos como:

$$w_i = 1, \forall i = 1, \dots, N$$

- ii. Cada clúster se subdivide en 2 subclústeres (V_1, V_2), iniciando desde el clúster final $U[N]$, recorriendo de arriba hacia abajo. Para cada uno de estos subclústeres se obtiene su varianza:

$$V_i = w^T V w, i = 1, 2 \quad (16)$$

Donde:

$$w = \frac{\text{diag}(V^{-1})}{\text{trace}(\text{diag}(V^{-1}))} \quad (17)$$

Teniendo presente que se está empleando la matriz cuasidiagonal, tiene sentido utilizar los pesos de asignación inversa con el objetivo de calcular la varianza para el subclúster. Esto debido a que la asignación de varianza inversa es óptima para una matriz de covarianza diagonal como fue indicado por Bechis (2020).

iii. Luego, se calcula un factor ponderador a partir de la nueva matriz de covarianzas:

$$a_1 = 1 - \frac{V_2}{V_1 + V_2}; a_2 = 1 - a_1 \quad (18)$$

iv. De esta manera, los 2 ponderadores obtenidos se incorporan en la determinación de los pesos finales al interior de cada subclúster, así los pesos de los activos en los subclústeres se actualizan respectivamente mediante:

$$w_1 = a_1 * w_1 \quad (19a)$$

$$w_2 = a_2 * w_2 \quad (19b)$$

v. Finalmente, se repiten los pasos ii al iv sobre los subclústeres V_1 y V_2 hasta que todos los activos tengan asignado su respectivo peso dentro del portafolio.

Vale la pena señalar que los pesos se asignan de arriba hacia abajo (*top-down*) en función de la varianza dentro del subclúster (Bechis, 2020), generando competencia de los activos intraclúster². Así, se evita la competencia con activos pertenecientes a los demás clústeres, lo cual es la principal ventaja del HRP, sobre

2 Se habla de competencia intraclúster porque el peso asignado al grupo debe ser distribuido entre los activos que lo conforman. De esta manera, un activo perteneciente al clúster 1 no competirá en participación con activos de otros clústeres.

otros algoritmos de asignación óptima. De igual manera, el algoritmo HRP garantiza que los pesos sean positivos, es decir, que $0 \leq w_1 \leq 1$, y que se alcance la distribución completa de recursos ($\sum_{i=1}^N w_i = 1$), lo cual es consistente con el enfoque clásico MV de optimización de portafolios.

2. Aplicación y resultados

2.1 Datos

Con el propósito de desarrollar una aplicación del modelo HRP, se seleccionaron siete activos del mercado estadounidense para el periodo comprendido entre marzo de 2010 (Mar-2010) y diciembre de 2021 (Dic-2021), lo cual permitirá disponer de información mensual que será obtenida a través de Yahoo Finance. La implementación se realizará en Python siguiendo el desarrollo propuesto por López de Prado (2016), Tatsat *et al.* (2020), Mercader (2021).

En la tabla 1 se presentan los activos seleccionados para conformar el portafolio, en donde para cada uno se obtuvo la información de 142 precios de cierre ajustados con periodicidad mensual. A partir de estos precios se calcularon los respectivos retornos logarítmicos y algunas estadísticas descriptivas.

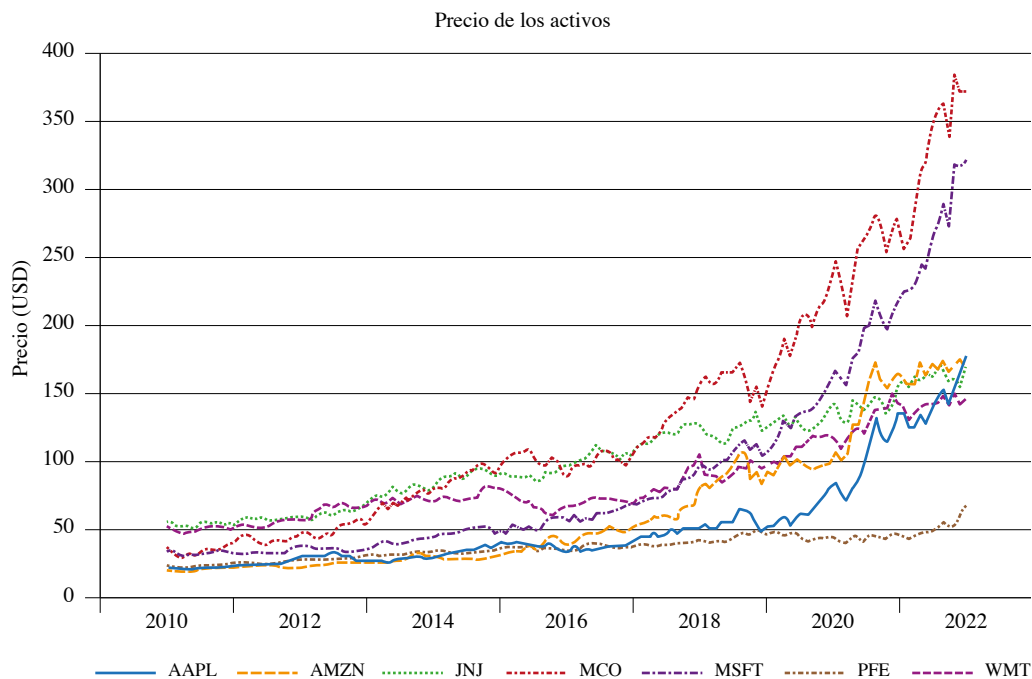
Tabla 1: Activos seleccionados

Ticker	Activo	Retorno promedio* (%)	Volatilidad* (%)	P-valor Test Jarque-Bera
AAPL	Apple	2,274	7,586	0,27
AMZN	Amazon	2,27	7,858	0,56
JNJ	Johnson y Johnson	0,9291	4,385	0,45
MCO	Moody's Corporation	1,933	7,264	0,08
MSFT	Microsoft	1,911	6,052	0,40
PFE	Pfizer	1,224	5,586	0,09
WMT	Walmart	0,8744	4,872	0,10

*Periodicidad mensual.

Fuente: elaboración propia.

Figura 2: Precio de cierre ajustado

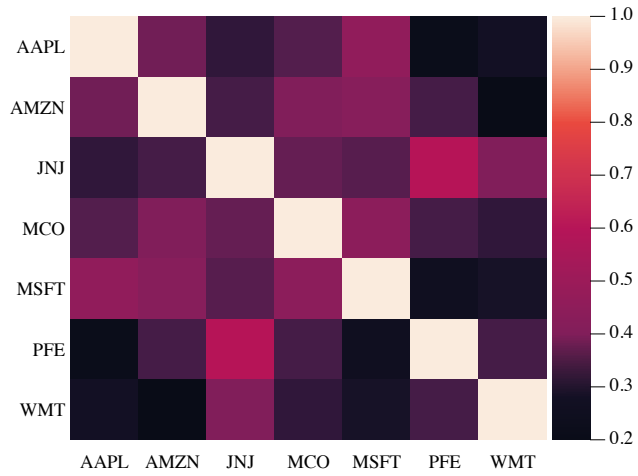


De acuerdo con la tabla 1, se destaca el caso de AAPL y AMZN que reportan los mayores retornos promedio y las más altas volatilidades del conjunto de activos, situación consistente con el enfoque de Markowitz. Por otra parte, los resultados del test de Jarque Bera evidencian que a un nivel del 5% no se puede rechazar la hipótesis nula acerca de que los retornos de los diferentes activos se distribuyen normalmente, esto es, no exhiben asimetría y curtosis que sea significativamente diferente de una distribución normal.

En este paso se construye la matriz de correlación con los retornos de los activos, destacando que el mayor valor se presenta entre JNJ y PFE con 0,54, y entre AAPL y MSFT con 0,47. Por su parte, el valor más bajo se da entre AMZN y WMT (0,17). Como se aprecia en la figura 3, la matriz de correlación presenta los activos ordenados alfabéticamente.

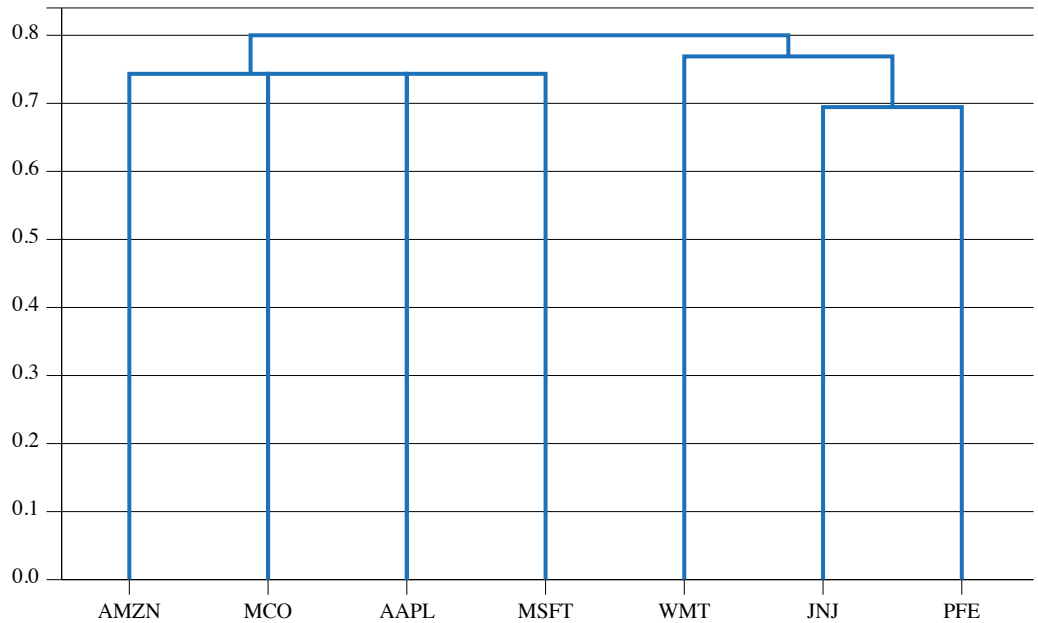
La primera parte del algoritmo de HRP consiste en jerarquizar los activos a partir de la información contenida en la matriz de correlación. En este paso, se emplea la distancia *single* mediante la cual se obtiene la matriz *linkage* que permite organizar los activos en un árbol agrupado, como se muestra en la figura 4.

Figura 3: Matriz de correlación



Fuente: elaboración propia.

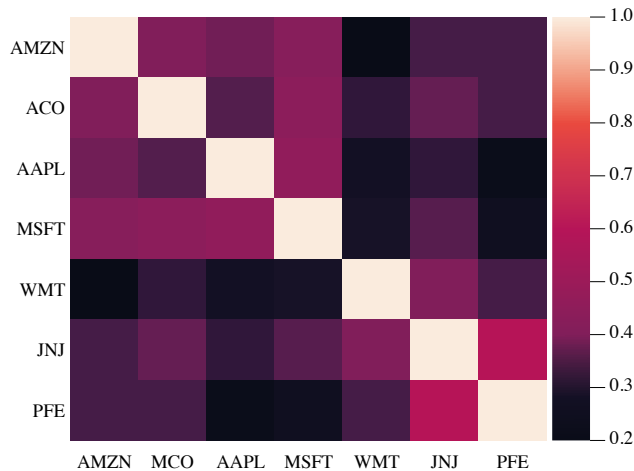
Figura 4: Dendrograma



Fuente: elaboración propia.

De acuerdo con el sector al cual pertenecen los activos seleccionados para el portafolio, se esperaba la conformación de clústeres de empresas tecnológicas y de la salud, lo cual es confirmado por el dendrograma en donde se refleja que los activos AAPL y MSFT conforman un clúster y que JNJ y PFE conforman otro grupo. Ahora bien, en la segunda parte del algoritmo se realiza el reordenamiento de la matriz de correlación conforme a lo encontrado en el primer paso (figura 5).

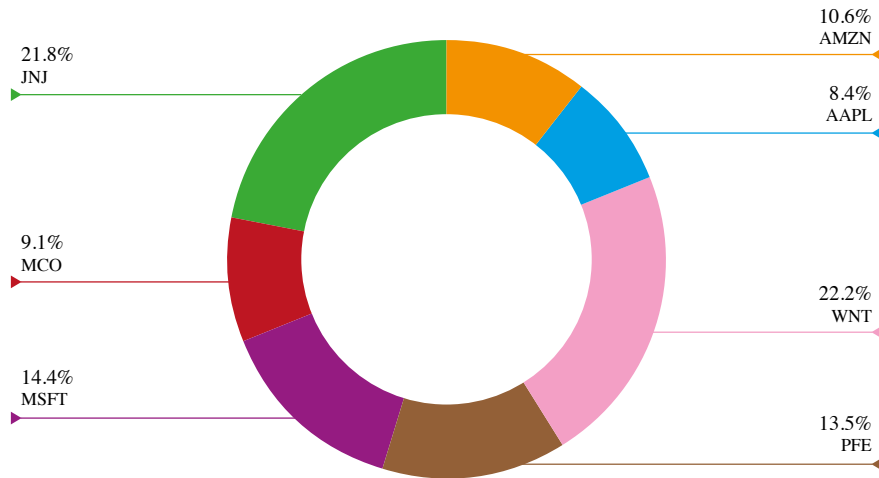
Figura 5: Matriz de correlación reordenada



Fuente: elaboración propia.

En el último paso del algoritmo HRP se realiza la asignación de pesos. Es importante mencionar que, dada la estructura de árbol jerárquico, se inicia con un peso igual a 1 correspondiente a todo el portafolio (parte alta del dendrograma). Seguido a esto, comienza la bifurcación en 2 clústeres, proceso que se realiza recursivamente de arriba hacia abajo (*top-down*), en donde finalmente a cada activo se le asigna el peso que se observa en la figura 6.

Figura 6: Distribución de pesos, HRP



Fuente: elaboración propia.

2.2 Análisis y comparación de resultados³

Con el propósito de comparar el resultado del modelo HRP, se estimaron dos modelos adicionales: media-varianza (MV)⁴ y paridad de riesgo (RP).

Tabla 2: Distribución de pesos, modelos HRP, MV y *risk parity* (RP)

Activo	HRP (%)	MV (%)	RP (%)
AAPL	8,387	5,267	8,436
AMZN	10,56	1,554	7,861
JNJ	21,83	32,07	25,24
MCO	9,146	0,17	9,199
MSFT	14,4	15,13	13,25
PFE	13,45	14,63	15,56
WMT	22,22	31,18	20,45

Fuente: elaboración propia.

³ El código de Python utilizado en esta aplicación se encuentra disponible en https://github.com/daragon66/Modelo-Paridad-de-Riesgo-Jerarquico/blob/main/C%C3%B3digo_Modelo_Paridad_de_Riesgo_Jer%C3%A1rquico_HRP.ipynb

⁴ La implementación del modelo MV se realiza con la restricción de pesos en corto.

Como se puede observar, el modelo de media-varianza (MV) asigna una mayor participación a los activos que presentan una menor volatilidad (JNJ y WMT) y, consecuentemente, reduce la participación de activos con alta volatilidad como AMZN y MCO, lo cual se traduce en una concentración cercana al 47% del total del portafolio en los dos activos pertenecientes al sector salud, situación que no se presenta en los otros dos portafolios en donde los activos en mención concentran entre un 35% (HRP) y 41% (RP).

Sumado a lo anterior, se observa que bajo el enfoque MV se concentra alrededor del 78% del portafolio en tres activos (JNJ, MSFT y WMT), en tanto que en los modelos HRP y RP este valor es cercano al 58%, situación que se atribuye al objetivo del MV (López de Prado, 2016): minimizar el riesgo del portafolio.

Respecto al modelo RP, se evidencian diferencias relativamente pequeñas en los pesos asignados frente al HRP. Se debe recordar que aquí la asignación de pesos obedece a la igualdad en la contribución al riesgo por cada activo. Así mismo, el resultado del modelo RP muestra a este portafolio como intermedio entre HRP y MV.

Igualmente, para cada uno de los tres modelos se obtuvo: retorno esperado, volatilidad, coeficiente de Sharpe y coeficiente de diversificación. Aquí, el modelo HRP es el que reporta un mayor rendimiento esperado y un mayor coeficiente de Sharpe; sin embargo, también presenta una mayor volatilidad frente a los otros dos modelos. A su vez, se evidencia un mayor coeficiente de diversificación para el modelo jerárquico, lo cual indica mejor distribución de los pesos entre los activos que conforman el portafolio.

Tabla 3: Desempeño de los portafolios

Portafolio	Retorno esperado (%)	Volatilidad (%)	Coeficiente de Sharpe	Coeficiente de diversificación
HRP	1,444	3,816	0,3784	1,518
MV	1,197	3,624	0,3304	1,433
RP	1,405	3,774	0,3723	1,51

Fuente: elaboración propia.

3. Conclusiones

El modelo de paridad de riesgo jerárquico (HRP) se apoya en técnicas de aprendizaje no supervisado para la optimización de portafolios, constituyéndolo como una alternativa al modelo clásico de media-varianza (V), por cuanto no se

requiere invertir la matriz de covarianzas de los retornos, lo cual evita situaciones relacionadas con la alta condición número en esta matriz.

La asignación óptima de los siete activos mediante el uso del modelo HRP evidencia que la distribución de estos en diferentes clústeres logra mejoras en términos del retorno esperado, así como del coeficiente de Sharpe frente a los resultados del modelo de MV. En menor medida, esta mejora también se presenta respecto del modelo RP. No obstante, es importante mencionar que, dados los activos contemplados para la implementación, el modelo HRP también obtuvo la mayor volatilidad.

Los resultados de la implementación muestran una mayor concentración bajo el enfoque MV frente a los modelos HRP y RP, lo cual es consistente con el objetivo de minimizar el riesgo del portafolio. Además, siguiendo a López de Prado (2016), esta concentración en algunos activos expone al portafolio MV a choques idiosincráticos.

Finalmente, estos resultados se deben enriquecer con la incorporación de otros modelos, medidas de desempeño, un conjunto más amplio de activos, además de un ejercicio dentro y fuera de muestra, de tal manera que se cuente con más información frente a las ventajas y desventajas del modelo HRP.

Referencias

- Bechis, L. (2020). Machine learning portfolio optimization: Hierarchical risk parity and modern portfolio theory (Tesis de maestría). Libera Università Internazionale degli Studi Sociali Guido Carli. http://tesi.luiss.it/28022/1/709261_BECHIS_LUCA.pdf
- Bailey, D. y López de Prado, M. (2012). The Sharpe coefficient efficient frontier. *Journal of Risk*, 15(2): 3-44. <https://doi.org/10.21314/JOR.2012.255>
- Black, F. y Litterman, R. (1992). Global portfolio optimization. *Financial Analysts Journal*, 48(5), 28-43. <https://doi.org/10.2469/faj.v48.n5.28>
- Clarke, R., De Silva, H. y Thorley, S. (2002). Portfolio constraints and the fundamental law of active management. *Financial Analysts Journal*, 58: 48-66. <https://doi.org/10.2469/faj.v58.n5.2468>
- Ledoit, O. y Wolf, M. (2004). A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 88(2), 365-411. [https://doi.org/10.1016/S0047-259X\(03\)00096-4](https://doi.org/10.1016/S0047-259X(03)00096-4)

- León, D., Aragón, A., Sandoval, J., Hernández, G., Arévalo, A. y Niño, J. (2017). Clustering algorithms for risk-adjusted portfolio construction. *Procedia Computer Science*, 108, 1334-1343. <https://doi.org/10.1016/j.procs.2017.05.185>
- López de Prado, M. (2016). Building diversified portfolios that outperform out of sample. *The Journal of Portfolio Management*, 42(4), 59-69. <https://doi.org/10.3905/jpm.2016.42.4.059>
- López de Prado, M. (2018). *Advances in financial machine learning*. John Wiley & Sons.
- López de Prado, M. (2020). *Machine learning for asset managers*. Cambridge University Press.
- Markowitz, H. (1952). Portfolio Selection. *Journal of Finance*, 7(1), 77-91.
- Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investments*. Wiley.
- Mercader, M. (2021). *Hierarchical Risk Parity: portfolio optimization*. Mathematics and Physics Engineering Final Project. Universitat Politècnica de Catalunya. <https://upcommons.upc.edu/bitstream/handle/2117/350200/TFG.pdf?sequence=1&isAllowed=y>
- Michaud, R. O. y Michaud, R. (2007). Estimation error and portfolio optimization: A Resampling Solution. Working paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2658657
- Raffinot, T. (2018). The hierarchical equal risk contribution portfolio. Working paper. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3237540.
- Sharpe, W. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance*, 19(3), 425-442. <https://doi.org/10.1111/j.1540-6261.1964.tb02865.x>
- Tatsat, H., Puri, S. y Lookabaugh, B. (2020). *Machine Learning and Data Science Blueprints for Finance*. O'Reilly Media.
- Vyas, A. (2019). The hierarchical risk parity algorithm: An introduction. Hudson and Thames Quantitative Research. Working paper. https://hudsonthames.org/?avia_forced_reroute=1