

Optimización de portafolio con cardinalidad usando algoritmos genéticos de población dual

Portfolio optimization with cardinality using dual population genetic algorithms

Sergio Iván Vanegas Gutiérrez*

* Maestría en Finanzas. Analista identificación y medición riesgo de mercado en Banco de Bogotá [sevanegas1@gmail.com]; [sergio.vanegas01@est.uexternado.edu.co]; [ORCID ID: 0000-0003-0106-9619].

Artículo recibido: 10 de diciembre de 2023

Aceptado: 28 de marzo de 2024

Para citar este artículo:

Vanegas Gutiérrez, S. I. (2024). Optimización de portafolio con cardinalidad usando algoritmos genéticos de población dual. *Odeon*, 26, pp. 95-126.

DOI: <https://doi.org/10.18601/17941113.n26.05>

Resumen

En algunos estudios se ha encontrado que, muchas veces, los algoritmos genéticos con una sola población para la solución de la optimización de portafolio con cardinalidad convergen lentamente y no obtienen los mejores resultados. Una manera de mejorar el desempeño de estos algoritmos ha sido incorporar una población adicional que actúe como buscador de máximos y mínimos locales; de esta manera, se aumenta la probabilidad de encontrar el óptimo global de la solución en un menor tiempo.

Este documento busca identificar el rendimiento en muestra y fuera de muestra de un portafolio de activos de renta variable con restricciones de cardinalidad usando algoritmos genéticos con una sola población y con población doble, estableciendo como universo el índice Dow Jones. Los resultados muestran que el desempeño puede verse afectado por los parámetros seleccionados para realizar la optimización, por lo que es importante tener en cuenta el error en la estimación de la media y la varianza del portafolio.

Palabras clave: optimización de portafolio; cardinalidad; algoritmos genéticos; portafolio de Markowitz; algoritmos evolutivos; población múltiple.

Clasificación JEL: C63, D81, G11.

Abstract

Some studies have found that genetic algorithms using a single population for portfolio optimization with cardinality constraints often converge slowly and do not achieve the best results. One way to improve the performance of these algorithms has been to incorporate an additional population that acts as a seeker of local maxima and minima; this increases the likelihood of finding the global optimum of the solution in a shorter time.

This document seeks to identify the in-sample and out-of-sample performance of an equity portfolio with cardinality constraints using genetic algorithms with a single population and with a double population, using the Dow Jones index as the universe. The results show that performance can be affected by the parameters selected for optimization, so it is important to consider the error in estimating the mean and variance of the portfolio.

Key words: Portfolio optimization; cardinality; genetic algorithms; markowitz portfolio; evolutionary algorithms; multiple population.

JEL classification: C63, D81, G11.

Introducción

Muchas investigaciones han hecho modificaciones del modelo de Markowitz (1952), incorporando restricciones prácticas a las que se enfrentan los gestores de portafolio. El problema de optimización de portafolio se vuelve complejo al solucionarlo con enfoques numéricos tradicionales cuando se incluyen restricciones que modelan el sentimiento del inversor y las fricciones que pueden ocurrir en el mercado. En la práctica, los gestores de portafolio tienen que operar bajo restricciones estrictas; una de ellas es la restricción de cardinalidad que, de acuerdo con Jimbo *et al.* (2017), ocurre cuando el inversionista o gestor solo puede invertir en K activos de un universo de N activos, para un valor K preespecificado. Como lo mencionan Ruiz y Suárez (2015), una de las razones principales para asumir una restricción de cardinalidad es que el limitar el número de activos en que el portafolio invierte hace la gestión del portafolio más simple. Adicionalmente, según Jalota y Thakur (2018), las restricciones de cardinalidad se usan para reducir el costo de transacción evitando un monto más alto al hacer una pequeña inversión en un número mayor de activos. Así mismo, de acuerdo con Maringer (2005) una característica importante en la selección de portafolio es que no es necesaria una selección amplia de activos para diversificar el riesgo, la diversificación de riesgo se puede lograr con un conjunto pequeño y con buenas selecciones de activos.

El inconveniente de incorporar las restricciones de cardinalidad en el modelo de Markowitz (1952) es que se vuelve complejo, ya que cambia a un problema de programación cuadrática entera mixta haciéndolo difícil de resolver por métodos numéricos tradicionales; por tal razón es necesario un método heurístico como el enfoque de los algoritmos genéticos.

Los motivos de usar algoritmos genéticos van más allá de la programación cuadrática entera mixta, según Gen y Cheng (2000) una propiedad importante es que los algoritmos evolucionarios son útiles cuando la función objetivo tiene varios óptimos locales. La característica principal de los algoritmos genéticos es la búsqueda global y multidireccional, en donde una población de soluciones potenciales se mantiene de generación en generación. Además, en la investigación de Sinha *et al.* (2015) se menciona que el enfoque población a población es benéfico en la exploración de soluciones para el portafolio óptimo, esta particularidad es fundamental ya que permite al algoritmo encontrar una solución global sin las restricciones de un mínimo retorno aceptable, proveyendo la mejor y más aceptable solución.

Sin embargo, estudios como el de Sabar y Song (2014) han encontrado que los algoritmos genéticos que utilizan una sola población para la solución de un portafolio con cardinalidad generalmente convergen lentamente, y esto puede evitar que obtengan buenos resultados. La solución para lo anterior ha sido incluir una población adicional con el fin de aumentar el espacio de búsqueda y encontrar el óptimo global de la solución en un menor tiempo. En la investigación de Sabar y Song (2014), los algoritmos genéticos de una sola población y los algoritmos genéticos con población dual han sido comparados en términos de eficiencia y resultados computacionales; pese lo anterior, estas metodologías de algoritmos genéticos con una sola población y población dual para portafolios con cardinalidad no han sido utilizadas en una instancia para la medición del rendimiento del portafolio.

Por esta razón, el propósito de esta investigación es estudiar el rendimiento en muestra y fuera de muestra de un portafolio de activos de renta variable con restricciones de cardinalidad usando algoritmos genéticos con población simple y población dual, y determinar cuál metodología ofrece mejores resultados. Para esto, se tendrán en cuenta las principales acciones que cotizan en el mercado de Estados Unidos, estableciendo como universo de acciones el índice Dow Jones, y se utilizarán los indicadores de Sharpe, Treynor, Sortino, Alfa de Jensen e *Information Ratio* para comparar el rendimiento del portafolio seleccionado.

1. Algoritmos genéticos

Los algoritmos genéticos son una optimización numérica inspirada en la selección natural y la genética natural, este método puede utilizarse en un rango amplio de problemas (Coley, 1999, p. 1). Un algoritmo genético permite que una población compuesta de muchos individuos evolucione bajo reglas específicas de selección a un estado que maximice o minimice la función objetivo (Haupt y Haupt, 2004, p. 22).

1.1. Operadores

Los principales operadores genéticos son tres: selección, intercambio genético (*crossover*) y mutación.

1.1.1. Selección

Este operador selecciona cromosomas en la población para la reproducción. Entre más adecuado sea el cromosoma más veces será su probabilidad de ser seleccionado para reproducirse.

1.1.2. Crossover

Este operador selecciona aleatoriamente una ubicación e intercambia la secuencia anterior y posterior de esa ubicación entre dos cromosomas para crear dos crías.

1.1.2.1. Crossover de límite promedio

El operador de *crossover* propuesto por Ling y Leung (2007) está dirigido a codificaciones de números reales. A este operador se le denomina *crossover* de límite promedio. Inicialmente, cuatro cromosomas son generados de dos padres seleccionados; después, las mejores dos crías en términos de valor de ajuste se seleccionarán para reemplazar a sus padres.

El operador propuesto por Ling y Leung (2007) utiliza dos operaciones. Cruce promedio y cruce con límite.

1.1.2.1.1. Crossover promedio. El cruce promedio se define de la siguiente manera:

$$\begin{aligned} O_{Sc1} &= \frac{(P_1 + P_2)}{2} \\ O_{Sc2} &= \frac{(P_{max} + P_{min})(1 - w_a) + (P_1 + P_2)w_a}{2} \end{aligned} \quad (1)$$

1.1.2.1.2. Crossover con límite. El cruce con límite se define:

$$\begin{aligned} O_{Sc3} &= P_{max}(1 - w_b) + \max(P_1 + P_2)w_b \\ O_{Sc4} &= P_{min}(1 - w_b) + \min(P_1 + P_2)w_b \end{aligned} \quad (2)$$

Donde:

$$O_{Sc}k = [O_{S_1}k, O_{S_2}k, \dots, O_{S_{no-vars}}k], k = 1, 2, 3, 4$$

$$P_i = [P_{i_1}, P_{i_2}, \dots, P_{i_j}, \dots, P_{i_{no-vars}}], i = 1, 2; j = 1, 2, \dots, no - vars$$

$$para_{min}^j \leq p_{ij} \leq para_{max}^j,$$

$$P_{max} = [para_{max}^1, para_{max}^2, \dots, para_{max}^{no-vars}]$$

$$P_{min} = [para_{min}^1, para_{min}^2, \dots, para_{min}^{no-vars}]$$

Los cromosomas P1 y P2 son los padres seleccionados, las variables Wa y Wb son pesos definidos por el usuario, Max (P1, P2) denota el vector con cada elemento obtenido tomando el máximo entre el padre 1 (P1) y el padre 2 (P2), igualmente pasa con el vector Min (P1, P2) el cual utiliza el mínimo. Osc representa las crías (Ling y Leung, 2007, p. 9).

1.1.3. Mutación

Este operador aleatoriamente cambia algunos caracteres en un cromosoma. La mutación puede ocurrir en cada posición de la cadena de caracteres con alguna probabilidad, usualmente muy pequeña (Mitchell, 1999, p. 8).

1.2. Estructura del algoritmo genético simple

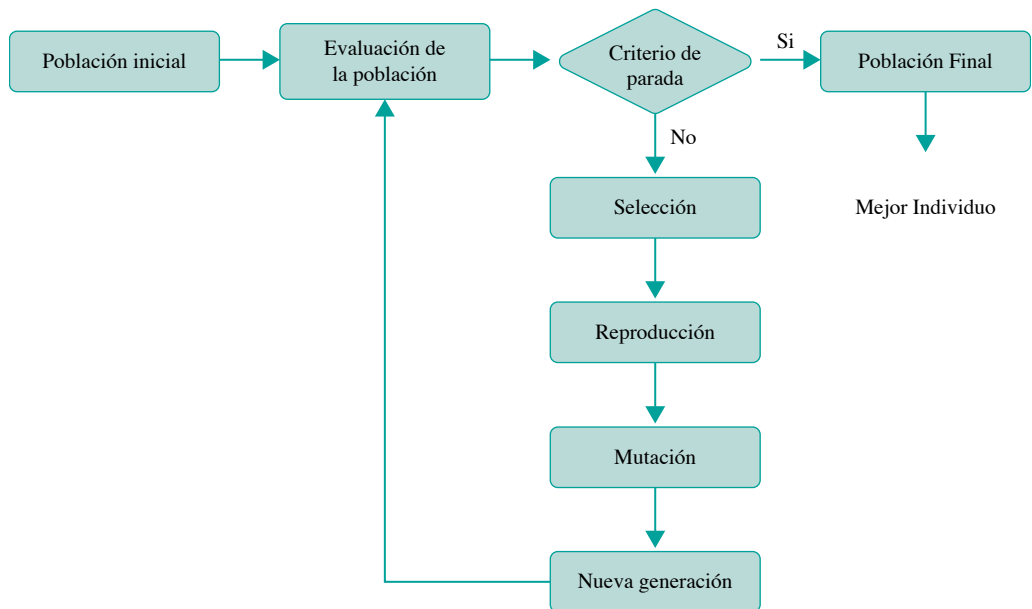
Según Mitchell (1999), un algoritmo genético simple funciona de la siguiente manera.

- i. Se comienza con una población aleatoriamente generada de n l-bits cromosomas (soluciones candidatas).
- ii. Se calcula la adecuación en la función objetivo de cada cromosoma en la población.
- iii. Repetir los siguientes pasos hasta que n crías hayan sido creadas.
 - a) Seleccionar un par de cromosomas progenitores de la población actual, la probabilidad de selección es una función creciente de adecuación. La selección se realiza por sustitución, es decir, el mismo cromosoma puede ser seleccionado más de una vez para convertirse en progenitor.
 - b) Con la probabilidad de *crossover* ocurre el intercambio de genes de par en un punto seleccionado aleatoriamente con probabilidad uniforme para

- formar dos crías. Si no ocurre el *crossover*, se forman dos crías exactamente iguales a los progenitores.
- c) Ocurre la mutación con probabilidad de mutación, y se colocan los cromosomas resultantes en la nueva población.
- iv. Se reemplaza la actual población con la población nueva.
 - v. Se va al paso 2.

Cada iteración es una generación y el conjunto entero de generaciones es conocido como un recorrido (figura 1).

Figura 1. Esquema algoritmo genético con una sola población



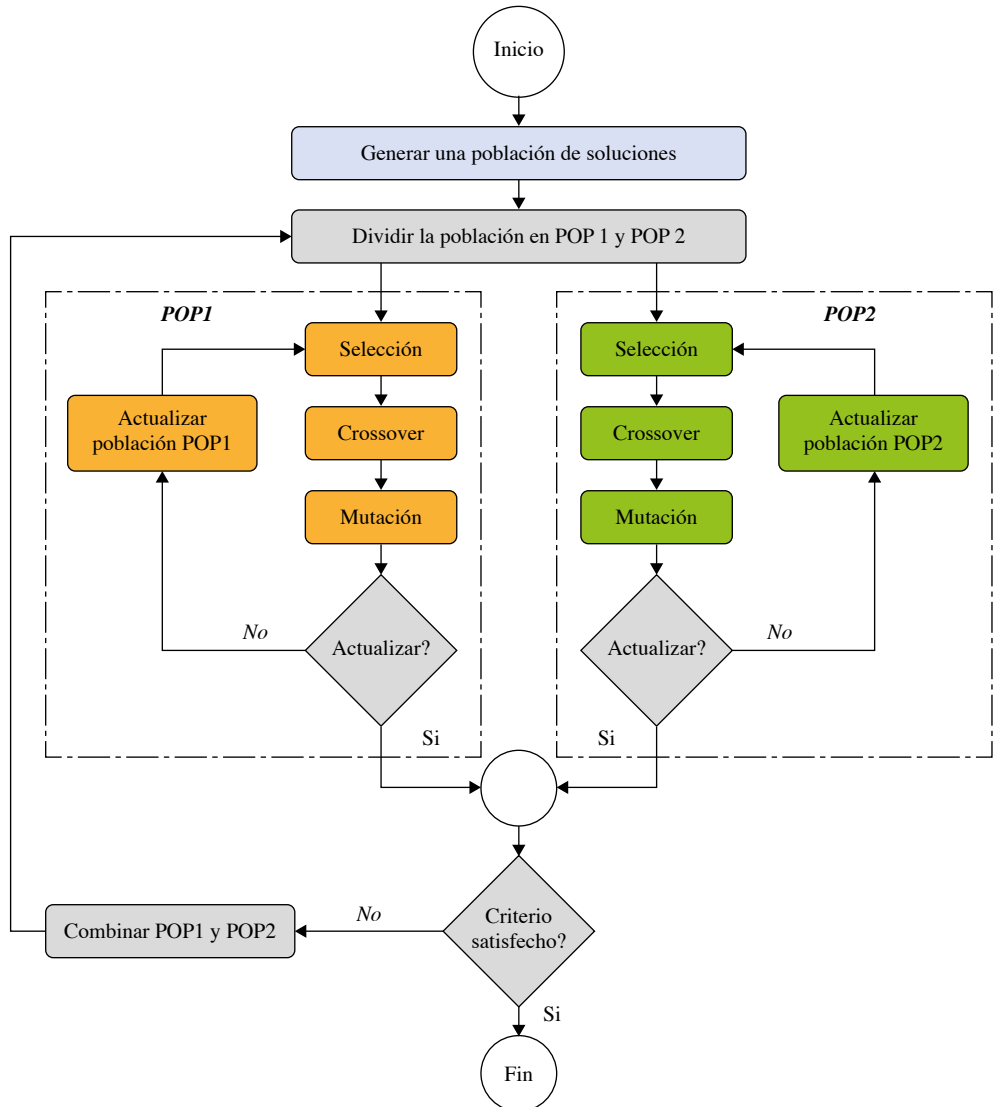
Fuente: Kramer (2017, p. 5).

1.3. Estructura del algoritmo genético con población dual

El modelo propuesto por Sabar y Song (2014) estimula el proceso de convergencia de los algoritmos genéticos y utiliza dos poblaciones. La primera población se encarga de la exploración del espacio de búsqueda, mientras que la segunda actúa como un operador de búsqueda local para mejorar el proceso

de convergencia. Las dos poblaciones intercambian individuos periódicamente para compartir soluciones prometedoras (figura 2).

Figura 2. Esquema algoritmo genético con población doble



Fuente: Sabar y Song (2014, p. 707).

POP1 es la población 1 y POP2 es la población 2.

1.4. Manejo de condiciones

De acuerdo con Michalewicz (1996), hay diferentes métodos de manejo de restricciones usados en los algoritmos genéticos. Este autor clasificó la mayoría de los métodos de manejos de restricciones evolucionarios en cuatro categorías.

- a) Métodos basados en preservar la factibilidad de las soluciones.
- b) Métodos basados en funciones de penalidad.
- c) Métodos que hacen distinción entre soluciones factibles y no factibles.
- d) Métodos híbridos.

Para esta investigación se utilizó el método propuesto por Deb (2000), que se encuentra clasificado como método basado en funciones de penalidad y métodos que hacen distinción entre soluciones factibles y no factibles.

1.4.1. Método de Deb

El método de Deb (2000) añade un término de penalidad a la función objetivo, sin embargo, el término de penalidad no se define como en los métodos convencionales. Este método usa el operador de método de selección por torneo, en donde dos soluciones se comparan al mismo tiempo y siguen los siguientes criterios:

- Cualquier solución factible es preferida frente a cualquier solución no factible.
- Entre dos soluciones factibles, la que tenga un mejor valor en la función objetivo es preferida.
- Entre dos soluciones no factibles, la que tenga la menor violación en las restricciones es preferida.
- Por lo anterior, la función de ajuste propuesta por Deb (2000) es la siguiente:

$$F(\vec{x}) = \begin{cases} f(\vec{x}) & \text{Si } g_j(\vec{x}) = 0 \quad \forall_j = 1, 2, \dots, m, \\ f_{max} + \sum_{j=1}^m & \text{De lo contrario} \end{cases} \quad (3)$$

- Donde las soluciones no factibles se comparan únicamente en la restricción de violaciones. El parámetro (f_{max}) es el valor de la función objetivo de la peor solución factible en la población. Esto hace que el ajuste de una solución no factible no solo dependa de la cantidad de la violación de la restricción, sino también de la población de soluciones.

2. Asignación de portafolio

Markowitz (1952) afirma que, para cualquier nivel dado de retorno esperado, un inversor racional seleccionaría el portafolio con varianza mínima entre un conjunto de todos los portafolios posibles.

El conjunto de todos los posibles portafolios que pueden ser construidos se llama el conjunto factible, los portafolios de mínima varianza se llaman portafolios de media varianza eficientes y el conjunto de todos los portafolios de media varianza eficientes, para diferentes niveles deseados de retorno esperado, se le conoce como la frontera eficiente (Fabozzi *et al.*, 2007, p. 22).

El modelo de Markowitz (1952) necesita los parámetros de los retornos esperados μ y la matriz de covarianza Σ . De acuerdo con Fabozzi *et al.* (2007, p. 25), Al seleccionar las ponderaciones del portafolio un inversor selecciona en los pares media-varianza disponibles. Para calcular las ponderaciones para un par posible, se debe establecer primero un retorno esperado objetivo, μ_0 . (p. 25).

De acuerdo con Markowitz (1952), el problema del inversor es un problema de minimización con restricciones, en donde se buscará minimizar la varianza del portafolio.

$$\begin{aligned} & \min w^T \Sigma w \\ \text{S.a} & \\ & \mu_0 = w^T \mu \\ & w^T l = 1, l^T = [1, 1, \dots, 1] \end{aligned} \quad (4)$$

La anterior formulación se refiere al problema de optimización clásica de media-varianza como una formulación de minimización de riesgo.

Otro tipo de formulación busca seleccionar un cierto nivel objetivo de riesgo para el portafolio, σ_0 , y luego maximizar el retorno esperado del portafolio.

$$\begin{aligned} & \max w^T \mu \\ \text{S.a} & \\ & w^T \Sigma w = \sigma_0 \\ & w^T l = 1, l^T = [1, 1, \dots, 1] \end{aligned} \quad (5)$$

Esta formulación se refiere como la maximización del retorno esperado del problema de optimización media-varianza.

2.1. Restricciones de cardinalidad

De acuerdo con Maringer (2005), las restricciones de cardinalidad son restricciones prácticas, en donde el inversor quiere mantener K activos diferentes en el portafolio de un universo de activos N .

Para utilizar las restricciones de cardinalidad, Chang *et al.* (2000) dan a conocer dos limitaciones en el modelo de Markowitz (1952): el supuesto de normalidad multivariada no es sostenible, ya que la distribución de los retornos de un activo individual tiende a exhibir alta probabilidad de los valores extremos, esto significa que se deben considerar momentos más altos. La otra limitación recae en la difícil aplicación de restricciones de enteros o restricciones de límites en la proporción del portafolio mantenido en un determinado activo.

El problema puede verse desde dos enfoques: el primero es utilizar algoritmos disponibles para resolver el problema de programación cuadrática entera mixta o reemplazar la medida de riesgo cuadrática de Markowitz (1952) por una medida que tenga una función lineal.

Para solucionar la limitación de las restricciones, Chang *et al.* (2000) utilizan métodos heurísticos y resuelve el problema utilizando la siguiente formulación:

$$\lambda \left[\sum_{i=1}^N \sum_{j=1}^N \sigma_{ij} \right] - (1 - \lambda) \left[\sum_{i=1}^N w_i \mu_i \right]$$

S.a

$$\begin{aligned} \sum_{i=1}^N w_i &= 1 \\ \sum_{i=1}^N z_i &= K \\ \epsilon z_i &\leq w_i \leq \delta_i z_i, \quad i = 1, \dots, N, \\ z_i &\in [0,1], \quad i = 1, \dots, N, \end{aligned} \tag{6}$$

Con esta formulación no es necesario establecer un mínimo retorno esperado, para ello se utiliza el parámetro lambda, donde lambda es el grado de aversión al riesgo entre 0 y 1, siendo 1 el grado más alto de aversión y 0 el menor grado de aversión. En esta especificación están incorporadas las restricciones de cardinalidad, presupuesto y cantidad en las ponderaciones.

3. Implementación de algoritmos genéticos en la optimización de portafolio

A continuación, se muestran las tablas resumen de los parámetros (tablas 1 y 2).

3.1. Datos económicos

Tabla 1. Datos económicos

Parámetro	Universo	Rendimiento mínimo esperado fuera de muestra	Rendimiento mínimo esperado en muestra	Cálculo Betas	Cálculo Alfa de Jensen	Cálculo Information Ratio
Dato	Índice Dow Jones	Rendimiento de mercado de los valores del Tesoro de EE. UU. con vencimiento constante de un mes	1 %	Índice Russell 3000	Dow Jones Industrial	Dow Jones Industrial
Fuente	Yahoo Finance	Datos económicos de la reserva federal	N/A	Yahoo Finance	Yahoo Finance	Yahoo Finance

Fuente: elaboración propia.

Para fines comparativos, el retorno mínimo esperado utilizado para contrastar el rendimiento en muestra fue del 1 %.

3.2. Parámetros generales

Tabla 2. Parámetros generales

Parámetro	Tiempo	Método Fuera de muestra	Longitud ventana de tiempo	Restricción ponderación	Lambda Chang et al. (2000)
Dato	01/12/2011– 31/12/2022	Demiguel y Gar- lappi, (2008)	120 meses	Min 5 %– Max 30 %	0,5
Fuente	N/A	Demiguel y Gar- lappi, (2008)	N/A	N/A	Chang et al. (2000)

Fuente: elaboración propia.

La función objetivo por optimizar es la propuesta por Chang *et al.* (2000). El parámetro que se usó para realizar las optimizaciones fue del 0,5 con el fin de maximizar los rendimientos y minimizar la varianza.

Inicialmente se calculan los rendimientos logarítmicos de las acciones, a partir de esta información se obtienen los retornos esperados, las volatilidades y las matrices de varianza y covarianza. Para cada uno de estos parámetros se realiza la respectiva optimización de portafolio asumiendo restricciones de cardinalidad. Los pasos que se siguen en la metodología de Demiguel y Garlappi (2008) son los siguientes:

- i. Se comienza en el periodo de tiempo $t-126$, en donde se estiman los parámetros requeridos para la optimización sobre la longitud seleccionada (120).
- ii. Con estos parámetros se resuelve el problema de optimización de portafolio para las metodologías de una sola población y población dual.
- iii. Se computa el resultado en el periodo $t + 1$ con la optimización obtenida, es decir, se obtiene el retorno del portafolio con las rentabilidades realizadas en el siguiente periodo, en este caso se utilizan las rentabilidades de las acciones del mes 121 y hallamos el rendimiento del portafolio con la optimización hecha para la ventana $M = 120$, de esta manera obtenemos el retorno del portafolio fuera de muestra.

La muestra rodante consiste en añadir los retornos del mes siguiente y quitar los retornos iniciales, esto mantiene la longitud de estimación fija. Para la segunda iteración, la ventana de estimación iría desde el mes 7 hasta el mes 126 y el

retorno fuera de muestra del portafolio sería para el mes 127. Este proceso se repite hasta llegar al final del conjunto de datos.

Una vez obtenidos los rendimientos de los portafolios, se calculan los indicadores de desempeño como Sharpe, Treynor, Alfa de Jensen, Information Ratio y Sortino. El cálculo de betas se realizó teniendo como referencia el índice Russell 3000. El índice de comparación para el Alfa de Jensen e Information Ratio fue el Dow Jones Industrial.

3.3. Parámetros generales algoritmos genéticos

Tabla 3. Parámetros generales algoritmos genéticos

Parámetro	Codificación GA	Individuos (Soluciones)	Operador crossover	Mutación
Dato	Números Reales	Vectores 1x30	Ling y Leung, (2007)	Flipping (Sivanandam y Deepa, 2008a)
Fuente	N/A	N/A	N/A	N/A

Fuente: elaboración propia.

3.4. Parámetros de algoritmos genéticos (simple 1 y dual 1)

La programación de los algoritmos genéticos se realizó en Python y se utilizaron las librerías: pytickersymbols, y finance, Pandas y Numpy.

3.4.1. Para el algoritmo genético con población simple 1

Tabla 4. Parámetros del algoritmo genético con población simple 1

Parámetros	Población
Tamaño de la población	100
Tasa de crossover	80 %
Tasa de mutación	5 %
Actualización de las poblaciones	Cada 100 evaluaciones de la función
Número de generaciones	800
Tiempo ejecución	2h 14min-2h 40 min

Fuente: elaboración propia.

3.4.2. Para el algoritmo genético con población dual 1

Tabla 5. Parámetros del algoritmo genético con población dual 1

Parámetros	Población1	Población2
Tamaño de la población	70	30
Tasa de crossover	80%	20 %
Tasa de mutación	5%	50 %
Actualización de las poblaciones	Cada 70 evaluaciones de la función.	Cada 30 evaluaciones de la función
Número de generaciones	800	
Tiempo de ejecución	2h 15min-2h 40 min	

Fuente: elaboración propia.

Las configuraciones simple 2, dual 2, simple 3 y dual 3 se encuentran en los anexos.

4. Resultados

4.1. Rendimiento índice Dow Jones (fuera de muestra)

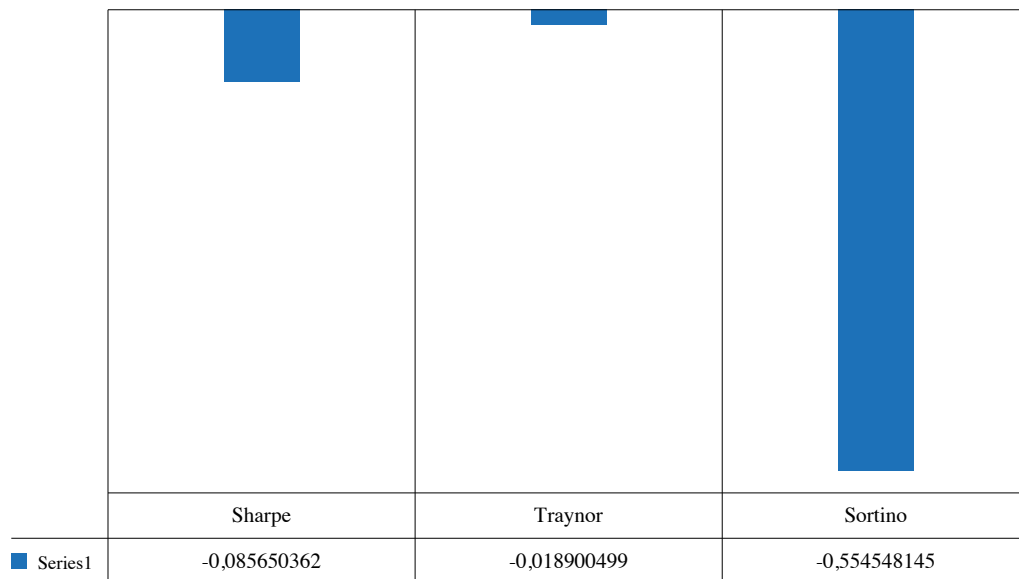
Tabla 6. Rendimiento Dow Jones

Fecha	Retorno Dow (%)	Volatilidad (%)	Beta (Russell)	Sharpe	Treynor	Valores del tesoro 1 mes (EU)	Sortino
1/07/2022	6,509	19,892	0,913	0,234	0,051	1,851%	1,472
1/08/2022	-4,148	19,980	0,907	-0,322	-0,071	2,276%	-2,031
1/09/2022	-9,253	20,043	0,909	-0,592	-0,131	2,613%	-3,798
1/10/2022	13,060	20,286	0,909	0,480	0,107	3,322%	3,031
1/11/2022	5,517	20,635	0,927	0,080	0,018	3,867%	0,508
1/12/2022	-4,260	20,684	0,928	-0,394	-0,088	3,899%	-2,510
Promedio	1,237	20,253	0,915	-0,086	-0,019	2,971%	-0,555

Fuente: elaboración propia.

En la tabla 6 se muestra que el resultado promedio para los indicadores Sharpe, Treynor y Sortino fue negativo para el periodo fuera de muestra seleccionado.

Figura 3. Indicadores fuera de muestra Dow Jones



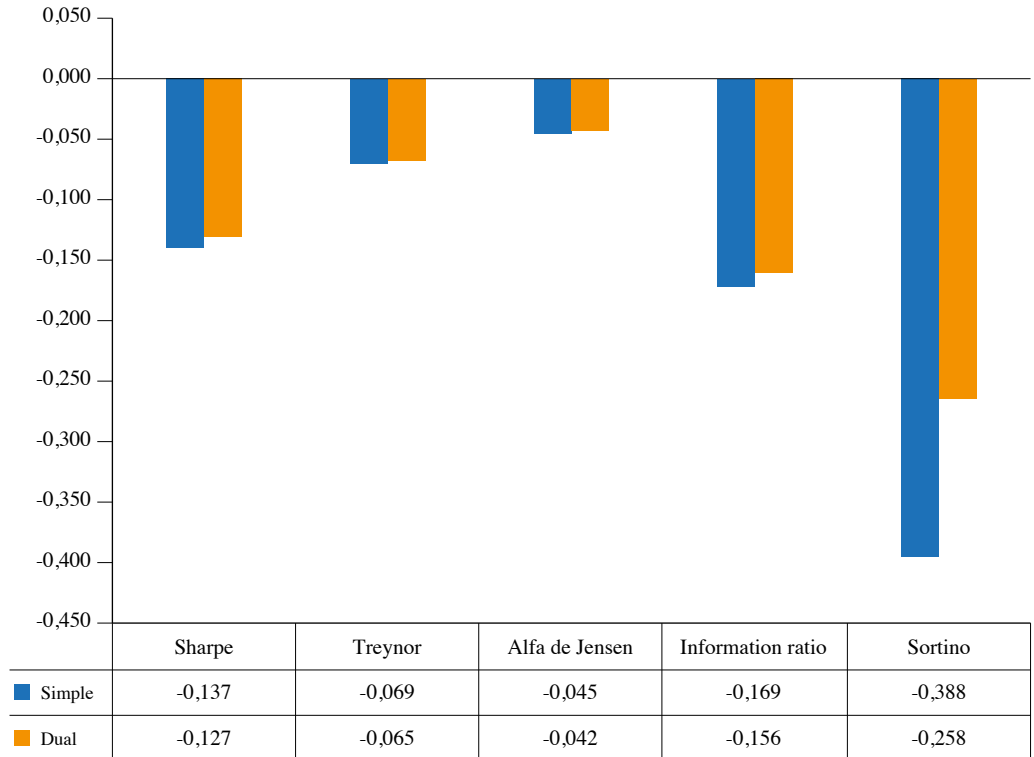
Fuente: elaboración propia.

Dado lo anterior, los resultados fuera de muestra que se presentarán con los algoritmos genéticos van a estar relacionados con el desempeño que tuvo el Dow Jones durante el periodo de estudio escogido.

4.2. Resultados fuera de muestra

4.2.1. Cardinalidad de 5

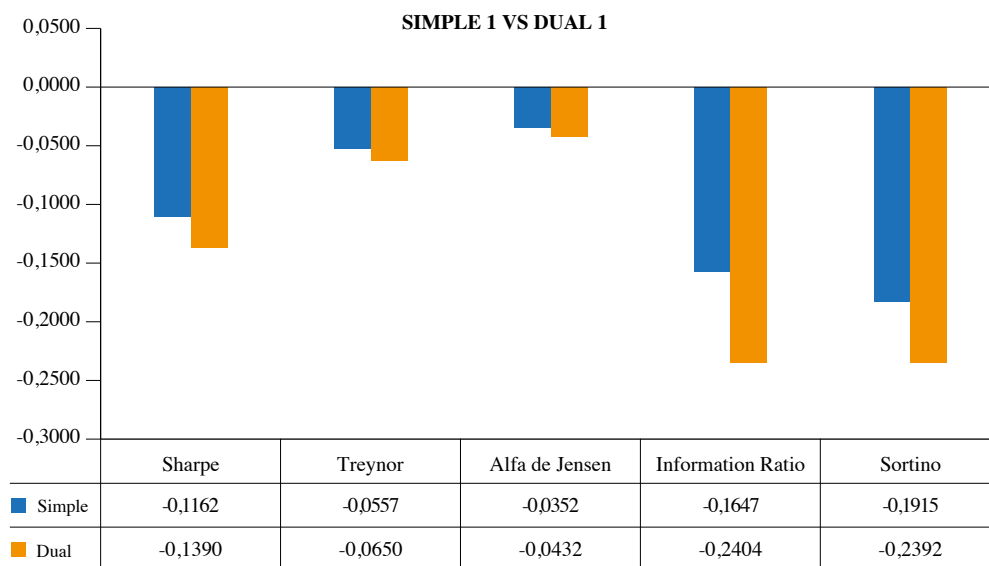
Figura 4. Indicadores fuera de muestra cardinalidad de 5
SIMPLE 1 VS DUAL 1



Fuente: elaboración propia.

4.2.2. Cardinalidad de 10

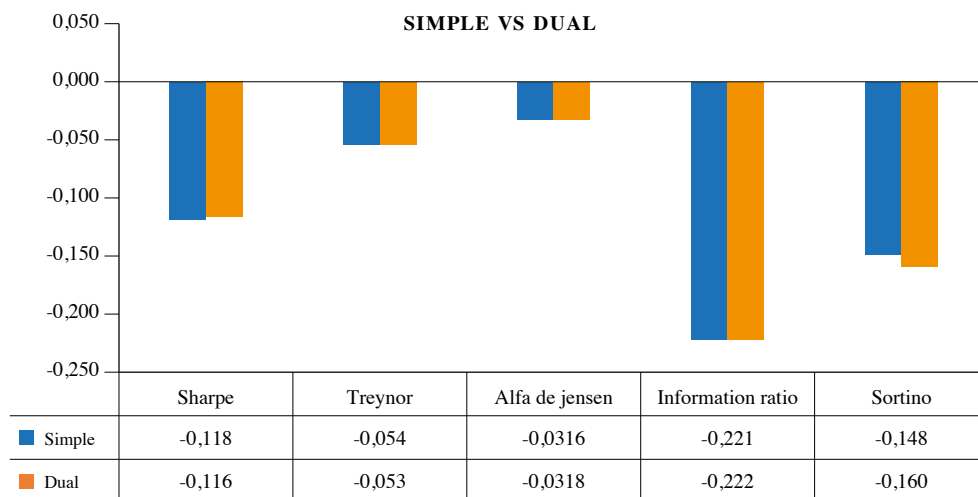
Figura 5. Indicadores fuera de muestra cardinalidad de 10



Fuente: elaboración propia.

4.2.3. Cardinalidad de 15

Figura 6. Indicadores fuera de muestra cardinalidad de 15



Fuente: elaboración propia.

El desempeño fuera de muestra de ambos algoritmos genéticos fue deficiente. Bajo diferentes niveles de cardinalidad, los retornos de las estrategias escogidas por los algoritmos fueron negativos, por lo que no superaban el rendimiento del mercado de los valores del Tesoro de Estados Unidos, con vencimiento constante de un mes. Esto en gran parte es resultado de utilizar los parámetros media y varianza como estimadores del mercado (Demiguel y Garlappi, 2008).

De igual forma, el resultado presentado por los algoritmos genéticos está altamente relacionado con el desempeño que tuvo el Dow Jones durante dicho periodo, como se mostró en el anexo.

Adicionalmente, se debe tener en cuenta que posiblemente el desempeño fuera de muestra de estos algoritmos puede mejorar sí se toma una mayor cantidad de datos, con el fin de que se tenga una mejor estimación de los parámetros del mercado.

4.3. Resultados en muestra por especificación

4.3.1. Cardinalidad de 5

Figura 7. Indicadores en muestra cardinalidad de 5 (todas las especificaciones)

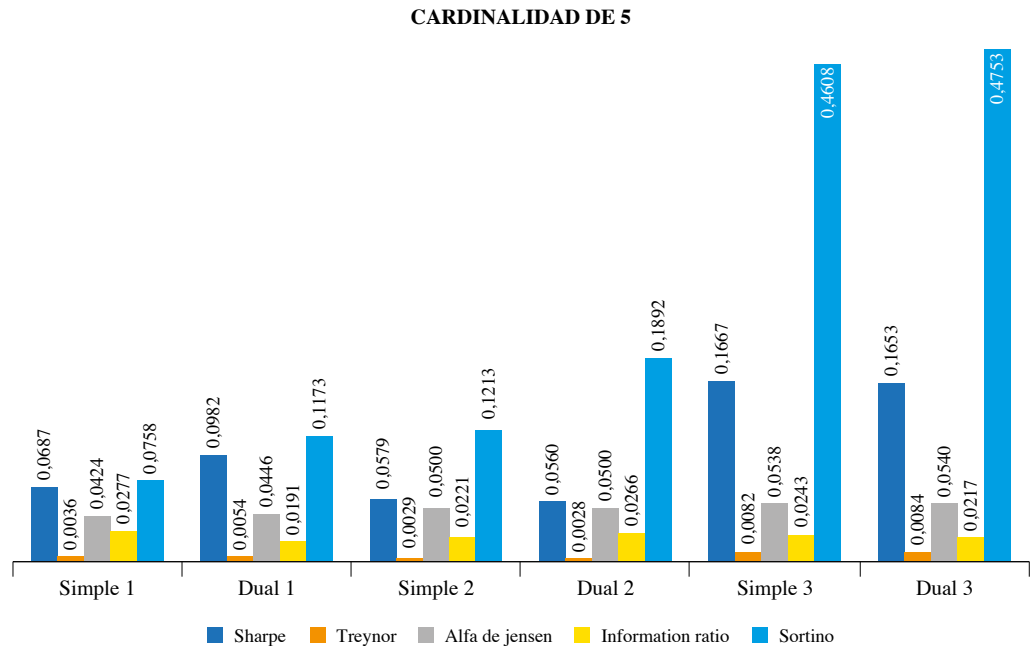


Tabla 7. Resumen mejores puntajes cardinalidad de 5

Indicador	Máximo	Mejor puntaje
Sharpe	0,1667	Simple 3
Treynor	0,0084	Dual 3
Alfa de Jensen	0,0540	Dual 3
Information Ratio	0,0277	Simple 1
Sortino	0,4753	Dual 3

Fuente: elaboración propia.

Para una cardinalidad de 5, la configuración que tuvo un mejor resultado fue la número 3 (ver anexos).

4.3.2. Cardinalidad de 10

Figura 8. Indicadores en muestra cardinalidad de 10 (todas las especificaciones)

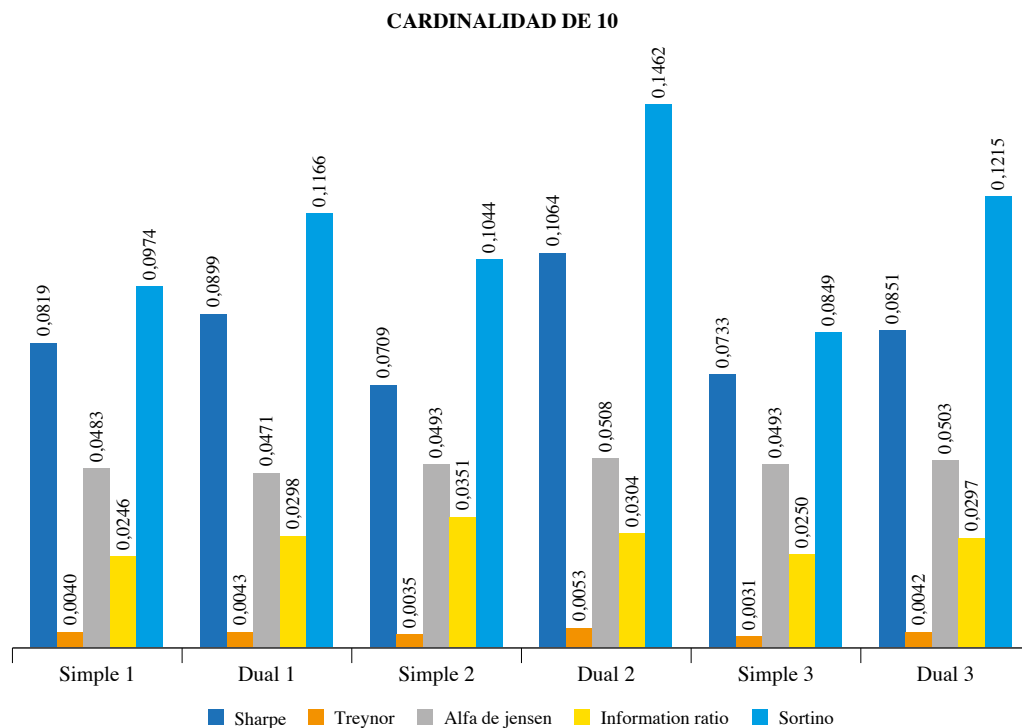


Tabla 8. Resumen mejores puntajes cardinalidad de 10

Indicador	Máximo	Mejor puntaje
Sharpe	0,1064	Dual 2
Treynor	0,0053	Dual 2
Alfa de Jensen	0,0508	Dual 2
Information Ratio	0,0351	Simple 2
Sortino	0,1462	Dual 2

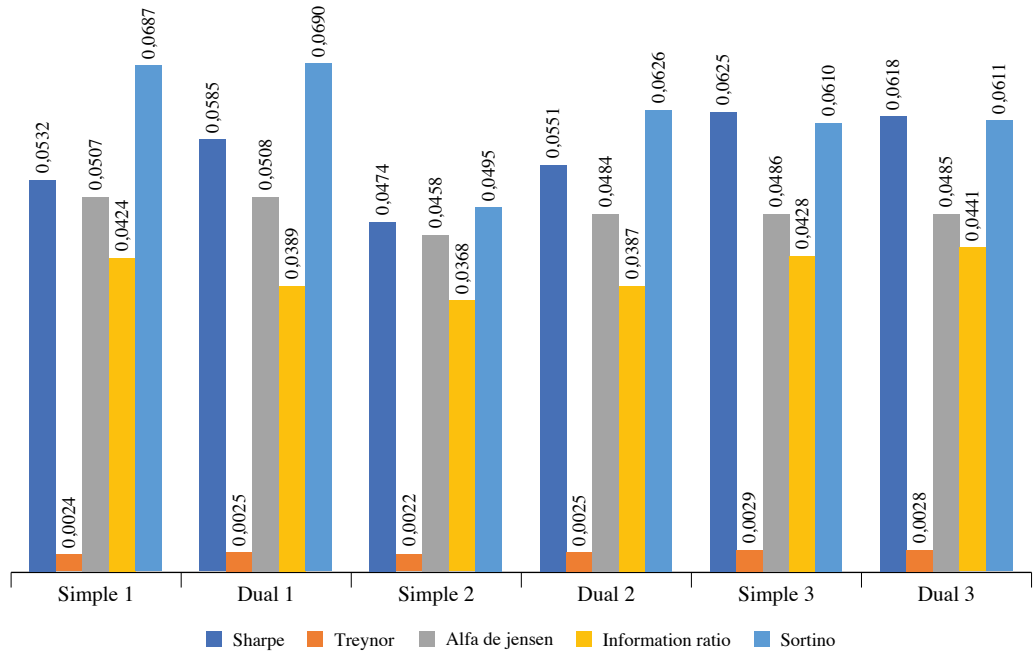
Fuente: elaboración propia.

Para una cardinalidad de 10, la configuración que tuvo un mejor resultado fue la número 2 (ver anexos).

4.3.3. Cardinalidad de 15

Figura 9. Indicadores en muestra cardinalidad de 15 (todas las especificaciones

CARDINALIDAD DE 15



Fuente: elaboración propia.

Tabla 9. Resumen mejores puntajes cardinalidad de 15

Indicador	Máximo	Mejor puntaje
Sharpe	0,0625	Simple 3
Treynor	0,0029	Simple 3
Alfa de Jensen	0,0508	Dual 1
Information Ratio	0,0441	Dual 3
Sortino	0,069	Dual 1

Fuente: elaboración propia.

Para una cardinalidad de 15, la configuración que tuvo un mejor resultado fue la número 3, seguida de la número 1.

4.3.4. Mejores puntajes para todas las cardinalidades y especificaciones

Tabla 10. Resumen mejores puntajes

Indicador	Máximo Total	Mejor puntaje total	Cardinalidad
Sharpe	0,1667	Simple 3	5
Treynor	0,0084	Dual 3	5
Alfa de Jensen	0,0540	Dual 3	5
Information Ratio	0,0441	Dual 3	15
Sortino	0,4753	Dual 3	5

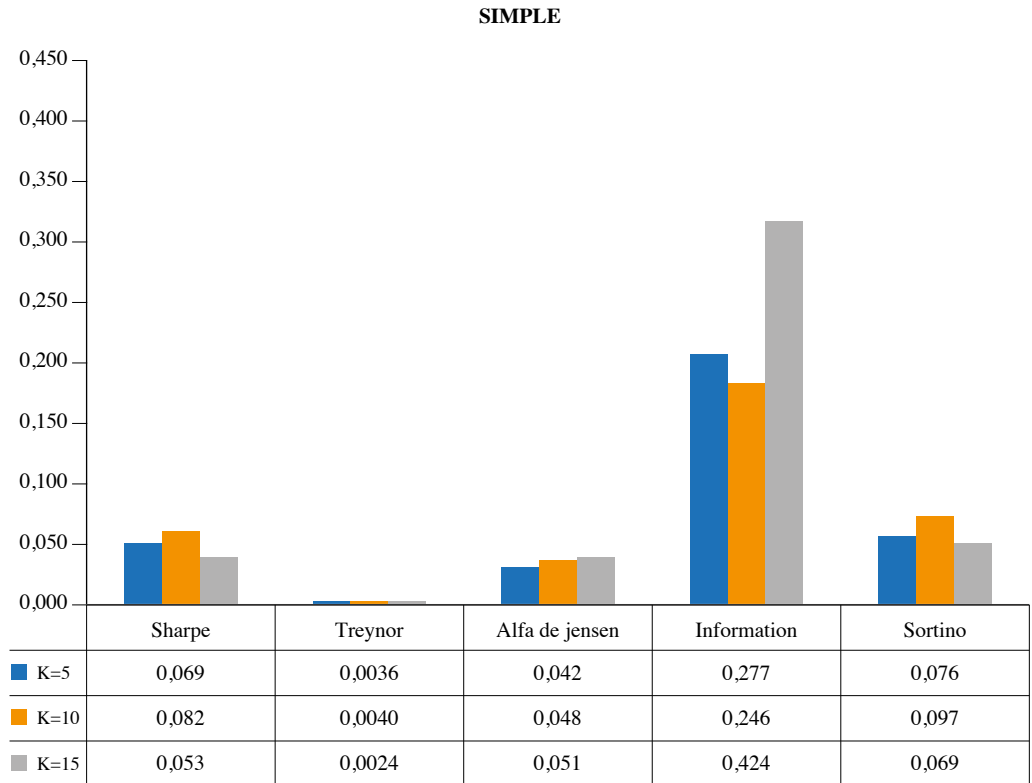
Fuente: elaboración propia.

Para todas las cardinalidades, la mayoría de los mejores resultados los tiene la cardinalidad de 5, a excepción del indicador Information ratio. Además, el algoritmo genético con población dual supera al simple en todos los indicadores a excepción del indicador de Sharpe.

4.4. Cardinalidad

4.4.1. Población simple

Figura 10. Cardinalidad población simple



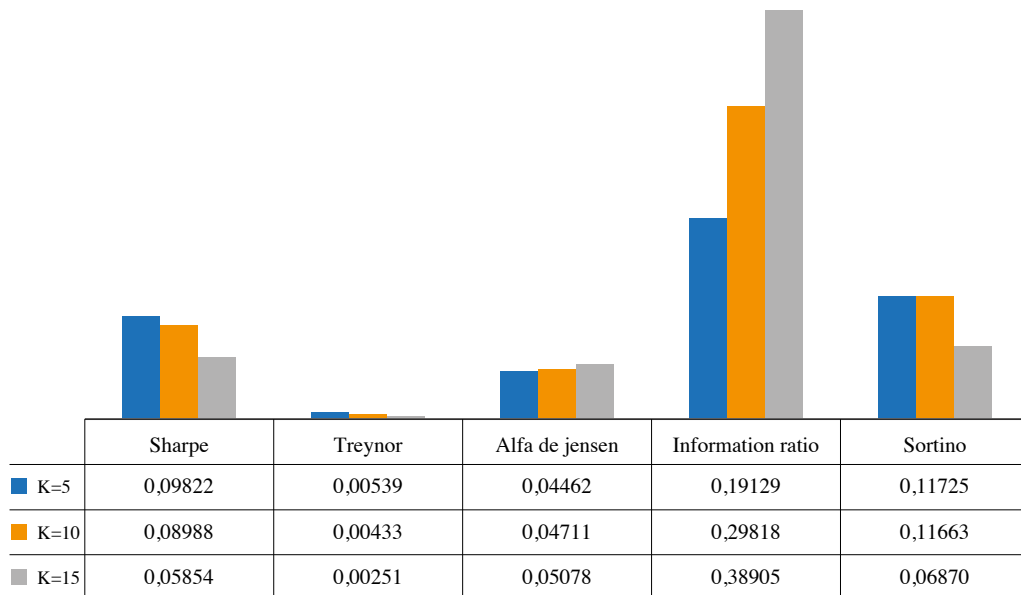
Fuente: elaboración propia.

En el algoritmo con población simple una mayor cardinalidad resultó en un menor retorno en el portafolio. La cardinalidad que obtuvo mejores resultados fue en un nivel de 10, la segunda cardinalidad que obtuvo mejores retornos fue de 5.

4.4.2. Población dual

Figura 11. Cardinalidad población dual

DUAL



Fuente: elaboración propia.

En el algoritmo con población dual, una menor cardinalidad resultó en un mejor desempeño de portafolio para la mayoría de los indicadores especialmente en Sharpe, Ratio y Sortino.

5. Conclusiones y extensión

- Los algoritmos genéticos son flexibles a cualquier tipo problema, en este caso para la optimización de portafolio se adapta bastante bien, los resultados obtenidos cumplen las restricciones establecidas en el modelo por lo que esta es una buena herramienta para resolver el problema de portafolio cuando se asumen restricciones adicionales.
- Al estudiar los resultados de las distintas ratios de rendimiento de portafolio utilizados para la metodología de población dual (Sabar y Song, 2014) y la metodología de población simple no es conveniente afirmar que la metodología propuesta por Sabar y Song (20214) presenta un mejor desempeño

fuera de muestra que el algoritmo genético con población simple, ya que el desempeño puede verse afectado por los parámetros seleccionados para realizar la optimización. Así mismo, los resultados fuera de muestra para ambos algoritmos fueron deficientes, esto es consecuencia de no tener en cuenta el error en la estimación de la media y la varianza. Por tal razón, es indispensable considerar métodos de estimación más robustos como los mencionados en la investigación de Demiguel y Garlappi (2008) para intentar mejorar el desempeño fuera de muestra.

De igual forma, los resultados negativos presentados están relacionados con el desempeño que tuvo el Dow Jones durante el periodo de tiempo seleccionado. Además, se pueden utilizar unas ventanas de tiempo más amplias con el fin de lograr una mejor estimación de los parámetros y obtener mejores resultados fuera de muestra.

En cuanto al desempeño en muestra, de acuerdo con los resultados obtenidos a nivel general, la metodología propuesta por Sabar y Song (2014), de algoritmos genéticos con poblaciones duales, presenta mejores resultados que los algoritmos genéticos con población simple, utilizando distintos parámetros de cardinalidad. Sin embargo, pese a que los algoritmos genéticos con poblaciones duales tienen una mayor probabilidad de obtener una mejor solución, los parámetros seleccionados para realizar la optimización pueden influir en el resultado al compararse con los algoritmos genéticos simples, como se muestra en los resultados y los anexos.

- El algoritmo genético con población dual tiene un mejor rendimiento en términos de Sharpe, Treynor y Sortino cuando la cardinalidad es inferior, mientras que el algoritmo con población simple tiene un mejor rendimiento con un nivel de cardinalidad igual a 10.

Esto prueba el concepto de Maringer (2005), quien menciona que no es necesario mantener portafolios con una alta cantidad de activos para una correcta gestión.

- Teniendo en cuenta los resultados de las diferentes configuraciones, el algoritmo con población dual superó más veces al algoritmo simple en los diferentes indicadores, sin embargo, pese a que los algoritmos genéticos con población dual tienen un mejor comportamiento que los algoritmos genéticos con población simple a nivel general, los resultados van a depender, en una gran proporción, de la selección de parámetros.

Para una cardinalidad de 5, una tasa de mutación que se encuentre entre el 2 y el 4 % y una tasa de *crossover* del 100 % pueden lograr mejores resultados

para la ratio de Sharpe, Treynor, Alpha de Jensen y Sortino, mientras que para el Information ratio una mutación del 5 % y una tasa de *crossover* del 80 % pueden ser mejores.

Para una cardinalidad de 10, una tasa de mutación del 1 y 60 % en las respectivas poblaciones, y una tasa de *crossover* del 90 % pueden lograr mejores resultados para todas las ratios.

Para una cardinalidad de 15, una tasa de mutación que se encuentre entre el 2 y el 4 %, y una tasa de *crossover* del 100 % pueden lograr mejores resultados para la ratio de Sharpe, Treynor e Information Ratio, mientras que para el alfa de Jensen y Sortino una mutación del 5 % y una tasa de *crossover* del 80 % pueden ser mejores.

Como extensión para esta investigación se propone comparar el esquema de población múltiple propuesto por Sabar *et al.* (2018) frente al esquema de población dual y simple. De igual forma, se puede plantear la utilización de métodos más robustos para la estimación de parámetros y usarlos con los algoritmos genéticos para la optimización del portafolio, y comparar su desempeño fuera de muestra, esto con el fin de obtener mejores resultados. Los métodos que se pueden usar son los expuestos por Demiguel y Garlappi (2008).

Referencias

- Amenc, N. y Le Sourd, V. (2003). *Portfolio Theory and Performance Analysis*. John Wiley & Sons Ltd.
- Arnone, S., Loraschi, A. y Tettamanzi, A. (1993). A genetic approach to portfolio selection. In *Neural Network World* (pp. 597-604). https://doi.org/10.1007/978-3-7091-7535-4_100
- Best, M. J. (2017). Quadratic programming with computer programs. In *Quadratic Programming with Computer Programs*. <https://doi.org/10.1201/9781315120881>
- Chang, T. J., Meade, N., Beasley, J. E. y Sharaiha, Y. M. (2000). Heuristics for cardinality constrained portfolio optimisation. *Computers and Operations Research*, 27(13), 1271-1302. [https://doi.org/10.1016/S0305-0548\(99\)00074-X](https://doi.org/10.1016/S0305-0548(99)00074-X)
- Christopherson, J. A., Cariño, D. Runge. y Ferson, W. E. (2009). *Portfolio performance measurement and benchmarking*. McGraw-Hill.

- Coley, D. (1999). *An introduction to genetic algorithms for scientists and engineers*. World Scientific.
- Deb, K. (2000). An efficient constraint handling method for genetic algorithms. *Computer Methods in Applied Mechanics and Engineering*, 186(2-4), 311-338. [https://doi.org/10.1016/S0045-7825\(99\)00389-8](https://doi.org/10.1016/S0045-7825(99)00389-8)
- DeMiguel, V., Garlappi, L. y Uppal, R. (2008). The Society for Financial Studies Optimal versus Naive Diversification: How Inefficient is the 1 / N Portfolio Strategy? *Review of Financial Studies*, 22(5), 1915-1953. <https://doi.org/10.1093/rfs/hhm075>
- Fabozzi, F. J., Kolm, P. N., Pachamanova, D. A. y Focardi, S. M. (2007). Robust portfolio optimization. *Journal of Portfolio Management*, 33(Issue 3). <https://doi.org/10.3905/jpm.2007.684751>
- Fama, E. F. (1965). The behavior of stock-market prices. *The Journal of Business*, 38(1), 34-105.
- Febrianti, W., Sidarto, K. A. y Sumarti, N. (2022). Solving constrained mean-variance portfolio optimization problems using spiral optimization algorithm. *International Journal of Financial Studies*, 11(1), 1. <https://doi.org/10.3390/ijfs11010001>
- Francis, J. y Kim, D. (2013). Modern portfolio theory foundations, analysis and new developments. In *Wiley Finance Series* (Vol. 4, Issue 1). John Wiley & Sons Ltd.
- FRED Economic Data (2023). *Market Yield on U.S. Treasury Securities at 1-Month Constant Maturity, Quoted on an Investment Basis*. Federal Reserve Economic Data. <https://fred.stlouisfed.org/series/DGS1MO>
- Gen, M. y Cheng, R. (2000). Genetic algorithms and engineering optimization. In *Wiley-interscience*. John Wiley & Sons Ltd.
- Guijarro, F. (2018). A similarity measure for the cardinality constrained frontier in the mean-variance optimization model. *Journal of the Operational Research Society*, 69(6), 928-945. <https://doi.org/10.1057/s41274-017-0276-6>
- Guijarro, F. y Tsinaslanidis, P. E. (2021). A surrogate similarity measure for the mean-variance frontier optimisation problem under bound and cardinality constraints. *Journal of the Operational Research Society*, 72(3), 564-579. <https://doi.org/10.1080/01605682.2019.1657367>

- Haupt, R. L. y Haupt, S. E. (2004). Practical genetic algorithms. In *Practical Genetic Algorithms*. <https://doi.org/10.1002/0471671746>
- Holland, J. H. (1992). *Adaptation in Natural and Artificial Systems. An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. MIT Press.
- Jalota, H. y Thakur, M. (2018). Genetic algorithm designed for solving portfolio optimization problems subjected to cardinality constraint. *International Journal of System Assurance Engineering and Management*, 9(1), 294-305. <https://doi.org/10.1007/s13198-017-0574-z>
- Jimbo, H. C., Ngongo, I. S., Andjiga, N. G., Suzuki, T. y Onana, C. A. (2017). Portfolio optimization under cardinality constraints: A comparative study. *Open Journal of Statistics*, 07(04), 731-742. <https://doi.org/10.4236/ojs.2017.74051>
- Kalayci, C. B., Polat, O. y Akbay, M. A. (2020). An efficient hybrid metaheuristic algorithm for cardinality constrained portfolio optimization. *Swarm and Evolutionary Computation*, 54, 100662. <https://doi.org/10.1016/j.swevo.2020.100662>
- Kolusheva, D. (2008). Out of Sample Performance of Asset Allocation Strategies. *SSRN Electronic Journal*.
- Konstantinou, C., Tzanetos, A. y Dounias, G. (2022). Cardinality constrained portfolio optimization with a hybrid scheme combining a Genetic Algorithm and Sonar Inspired Optimization. *Operational Research*, 22(3), 2465-2487. <https://doi.org/10.1007/s12351-020-00614-1>
- Kramer, O. (2017). *Studies in Computational Intelligence 679 Genetic Algorithm Essentials*. Springer-Berlin.
- Leyffer, S. y Lee, J. (2012). Mixed Integer Nonlinear Programming, MINLP. In *SpringerReference*. https://doi.org/10.1007/springerreference_72472
- Ling, S. H. y Leung, F. H. F. (2007). An improved genetic algorithm with average-bound crossover and wavelet mutation operations. *Soft Computing*, 11(1), 7-31. <https://doi.org/10.1007/s00500-006-0049-7>
- Loraschi, A., Tomassini, M., Tettamanzi, A. y Verda, P. (1995). Distributed genetic algorithms with an application to portfolio selection problems. *Artificial Neural Nets*

- and Genetic Algorithms* (Conference paper), 384-387. https://doi.org/10.1007/978-3-7091-7535-4_100
- Maringer, D. G. (2005). Portfolio management with heuristic optimization. In Springer (Ed.), *Portfolio Management with Heuristic Optimization* (Vol. 00008). Springer. <https://doi.org/10.1007/b136219>
- Markowitz, H. (1952). Portfolio Selection. *Journal of Finance*, 7(1), 77-91. <https://doi.org/10.2307/2975974>
- Michalewicz, Z. (1996). Evolutionary algorithms for constrained parameter optimization problems. *Evolutionary Computation*, 4(1), 1-32. <https://doi.org/10.1162/evco.1996.4.1.1>
- Mitchell, M. (1999). *An Introduction to Genetic Algorithms*. The MIT press.
- Moral-Escudero, R., Ruiz-Torrubiano, R. y Suárez, A. (2006). Selection of optimal investment portfolios with cardinality constraints. *2006 IEEE Congress on Evolutionary Computation, CEC 2006*, 2382-2388. <https://doi.org/10.1109/cec.2006.1688603>
- Ruiz-Torrubiano, R. y Suárez, A. (2010). Hybrid approaches and dimensionality reduction for portfolio selection with cardinality constraints. *IEEE Computational Intelligence Magazine*, 5(2), 92-107. <https://doi.org/10.1109/MCI.2010.936308>
- Ruiz-Torrubiano, R. y Suárez, A. (2015). A memetic algorithm for cardinality-constrained portfolio optimization with transaction costs. *Applied Soft Computing Journal*, 36, 125-142. <https://doi.org/10.1016/j.asoc.2015.06.053>
- Sabar, N. R. y Song, A. (2014). Dual population genetic algorithm for the cardinality constrained portfolio selection problem. In G. Dick, W. N. Browne, P. Whigham, M. Zhang, L. T. Bui, H. Ishibuchi, ... K. Tang (Eds.), *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 8886, pp. 703-712). Springer International Publishing. https://doi.org/10.1007/978-3-319-13563-2_59
- Sabar, N. R., Turkey, A., Leenders, M. y Song, A. (2018). Multi-population genetic algorithm for cardinality constrained portfolio selection problems. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 10860 LNCS, 129-140. https://doi.org/10.1007/978-3-319-93698-7_10

- Sadjadi, S. J., Gharakhani, M. y Safari, E. (2012). Robust optimization framework for cardinality constrained portfolio problem. *Applied Soft Computing Journal*, 12(1), 91-99. <https://doi.org/10.1016/j.asoc.2011.09.006>
- Sinha, P., Chandwani, A. y Sinha, T. (2015). Algorithm of construction of optimum portfolio of stocks using genetic algorithm. *International Journal of System Assurance Engineering and Management*, 6(4), 447-465. <https://doi.org/10.1007/s13198-014-0293-7>
- Sivanandam, S. N. y Deepa, S. N. (2008a). *Genetic Algorithms BT–Introduction to Genetic Algorithms* (S. N. Sivanandam & S. N. Deepa, Eds, pp. 15-37). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-73190-0_2
- Sivanandam, S. N. y Deepa, S. N. (2008b). *Introduction to Genetic Algorithms*. Springer. <https://doi.org/10.1007/978-3-540-73190-0>
- Soleimani, H., Golmakani, H. R. y Salimi, M. H. (2009). Markowitz-based portfolio selection with minimum transaction lots, cardinality constraints and regarding sector capitalization using genetic algorithm. *Expert Systems with Applications*, 36(3 part 1), 5058-5063. <https://doi.org/10.1016/j.eswa.2008.06.007>
- Woodside-Oriakhi, M., Lucas, C. y Beasley, J. E. (2011). Heuristic algorithms for the cardinality constrained efficient frontier. *European Journal of Operational Research*, 213(3), 538-550. <https://doi.org/10.1016/j.ejor.2011.03.030>
- Yaman, I. y Erbay Dalkılıç, T. (2021). A hybrid approach to cardinality constraint portfolio selection problem based on nonlinear neural network and genetic algorithm. *Expert Systems with Applications*, 169(December 2020), 114517. <https://doi.org/10.1016/j.eswa.2020.114517>

Anexos

Anexo A. Configuración algoritmos genéticos (simple 2 y dual 2)

Tabla 1. Parámetros del algoritmo genético con población simple 2

Parámetros	Población
Tamaño de la población	100

Parámetros	Población
Tasa de crossover	90 %
Tasa de mutación	1 %
Actualización de las poblaciones	Cada 100 evaluaciones de la función
Número de generaciones	800

Fuente: elaboración propia.

Tabla 2. Parámetros del algoritmo genético con población dual 2

Parámetros	Población1	Población2
Tamaño de la población	80	20
Tasa de crossover	90 %	20 %
Tasa de mutación	1 %	60 %
Actualización de las poblaciones	Cada 70 evaluaciones de la función	Cada 30 evaluaciones de la función
Número de generaciones	800	

Fuente: elaboración propia.

Anexo B. Configuración algoritmos genéticos (simple 3 y dual 3)

Tabla 3. Parámetros del algoritmo genético con población simple 3

Parámetros	Población
Tamaño de la población	100
Tasa de crossover	100 %
Tasa de mutación	3 %
Actualización de las poblaciones	Cada 100 evaluaciones de la función.
Número de generaciones	800

Fuente: elaboración propia.

Tabla 4. Parámetros del algoritmo genético con población dual 3

Parámetros	Población1	Población2
Tamaño de la población	80	20

Parámetros	Población1	Población2
Tasa de crossover	100 %	20 %
Tasa de mutación	3 %	70 %
Actualización de las poblaciones	Cada 70 evaluaciones de la función.	Cada 30 evaluaciones de la función.
Número de generaciones	800	

Fuente: elaboración propia.