

Selección óptima de cartera de inversión de stocks de S&P 500: Uso de algoritmos genéticos

Marco Adrián Guido Torres, Volodymyr Ponomaryov

Instituto Politécnico Nacional, SEPI de ESIME Culhuacan,
Ciudad de México, México

marco@guidotorres.mx, vponomar@ipn.mx

Resumen. El primer modelo de optimización de carteras de inversión fue propuesto por el doctor Markowitz en 1952 y es conocido como modelo de varianza promedio, es usado para encontrar la asignación optimizada de recursos entre diferentes activos financieros, obteniendo retornos de inversión mayores y menores riesgos de inversión. Sin embargo, el modelo de Markowitz no consideraba algunas limitaciones prácticas de los mecanismos de los mercados disponibles, entre ellos: I) costo de transacción, cantidad mínima de reasignación de recursos. Sin estas consideraciones, la funcionalidad del modelo se ve limitada, y diferentes enfoques se han utilizado para intentar resolver el problema, en este artículo se ataca usando algoritmos genéticos.

Palabras clave: portfolio, mean-variance model, genetic algorithms.

Optimal Selection of S&P 500 Stock Investment Portfolio: Use of Genetic Algorithms

Abstract. The first model of optimization of investment portfolios was proposed by Dr. Markowitz in 1952 and is known as the average variance model, it is used to find the optimized allocation of resources between different financial assets, obtaining greater investment returns and lower investment risks. However, the Markowitz model did not consider some practical limitations of the available market mechanisms, including: I) transaction cost, minimum amount of reallocation of resources. Without these considerations, the functionality of the model is limited, and different approaches have been used to try to solve the problem, in this article it is attacked using genetic algorithms.

Keywords: portfolio, mean-variance model, genetic algorithms.

1. Introducción

La forma óptima de selección de portafolios propuesta inicialmente por Markowitz [1] en 1952, utilizaba un enfoque de análisis estadístico clásico, para encontrar la

cartera de inversión óptima, pero no consideraba limitaciones prácticas relacionadas al costo de recolocación de recursos y de la cantidad mínima que se puede reasignar. Si se consideran estos factores, la complejidad del problema aumenta y sus restricciones cambian considerablemente, lo que se suma a un problema que ya es de complejidad NP, como se demuestra en [4].

Múltiples algoritmos heurísticos han demostrado un buen desempeño al atacar este tipo de problemas y [5], [6] y [7] lo hacen usando algoritmos genéticos, reinforcement learning y redes neuronales respectivamente, obteniendo buenos resultados que a su vez.

En este artículo se usa un algoritmo genético para extender el modelo de Markowitz y se aplica a los stocks que componen el índice S&P 500, NASDAQ y FTSE y se compara con el modelo estándar.

2. Descripción del problema

Un portafolio (o cartera) es un grupo de activos financieros como acciones, bonos y fondos en lo general, los portafolios son administrados por profesionales financieros, y su propósito es reconstruir continuamente un portafolio de inversión de acuerdo con el riesgo asociado de los activos incluidos y los objetivos de corto y largo plazo.

Para determinar una asignación de recursos adecuada entre los activos incluidos en un portafolio, se debe considerar el riesgo individual de cada activo y el riesgo asociado equivalente del grupo seleccionado, el portafolio elegido debe ser construido considerando una compensación entre el riesgo y el retorno de inversión asociado. Un portafolio será mejor que otro si tiene una mejor expectativa de ganancia y un menor riesgo asociado.

La teoría moderna de portafolios asume que los inversionistas son muy renuentes al riesgo; si dos portafolios ofrecen el mismo retorno esperado de inversión, típicamente se seleccionará el que tenga un menor riesgo, sin embargo, para que exista un motivo para arriesgarse por un portafolio ligeramente más riesgoso, las expectativas de retorno tienen que ser considerablemente mayores.

Además, si la correlación entre un activo con otro es alta, entonces el riesgo de que ocurran pérdidas generalizadas es mayor, por lo que es importante seleccionar activos con un bajo nivel de correlación para mantener el nivel de volatilidad controlada.

3. Modelo de Markowitz

El modelo de Markowitz se puede resumir a partir de las ecuaciones:

1. El retorno esperado de cada portafolio se define como:

$$E(R_p) = \sum_i w_i E(R_i) \quad (1)$$

$$\sum_i^n w_i = 1 \quad (2)$$

Donde:

- R_p representa el retorno del portafolio completo.
- R_i es el retorno específico de cada activo.
- w_i es la importancia relativa de cada activo. relativo al conjunto de recursos disponibles.
- n representa la cantidad de activos disponibles.

2. La volatilidad del portafolio de interés se define como:

$$\sigma_p^2 = \sum_i \sum_j w_i w_j \sigma_{ij} \quad (3)$$

Donde:

- σ_{ij} es la covarianza del retorno esperado de los activos,
 - σ_p^2 es la volatilidad y es proporcional al riesgo esperado del portafolio.
3. Teniendo en cuenta que (1) representa el retorno esperado y (3) el riesgo asociado de un portafolio específico, se puede plantear (4) como el modelo que se debe maximizar para obtener el máximo rendimiento con el mínimo riesgo posible:

$$\begin{aligned} & \text{maximizar} \\ & (1 - \lambda) \sum_i^n w_i R_i - \lambda \sum_i^n \sum_j^n w_i w_j \sigma_{ij} \end{aligned} \quad (4)$$

Donde: λ representa la aversión al riesgo deseada y está limitada por $0 < \lambda < 1$.

4. Para maximizar el retorno de inversión considerando los costos de transacciones, es necesario tomar en cuenta el costo de reasignación de recursos y la cantidad de recursos recolocados:

$$\begin{aligned} & \text{maximizar} \quad ((1 - \lambda) \sum_i^n w_i R_i - \\ & \lambda \sum_i^n \sum_j^n w_i w_j \sigma_{ij} - \sum_i^n c \Delta w_i) \end{aligned} \quad (5)$$

Donde:

- c representa el costo de transacción y lambda,
- Δw_i representa el peso reasignado del índice específico.

A partir de este modelo modificado, se puede observar que la función de evaluación del desempeño del portafolio debe estar dado por la ecuación 5.

4. Algoritmos genéticos

Los algoritmos genéticos están inspirados en la selección natural, y son usados para optimizar y buscar soluciones de problemas a través de operadores bioinspirados, mutación, reproducción, selección, etcétera.

A continuación se presentan las características específicas del algoritmo usado.

4.1. Generación de población inicial

Al representar el peso de una señal, la cual se trata como un gen y al considerar todos los genes como un vector de pesos, un vector compuesto de pesos constituye un cromosoma que representa un portafolio.

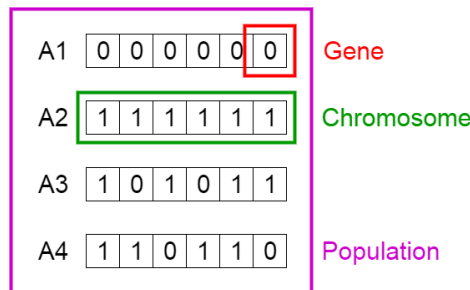


Fig. 1. Representación de la población inicial.

Como población inicial para este algoritmo se presentan 100 cromosomas de longitud variable, dependiendo de la cantidad de stocks que compongan el índice analizado, para S&P 500 se usan 500 genes, como se representa en la figura 1.

4.2. Evaluación

Para determinar una calificación del desempeño de cada individuo de la población creada, se usa una función adecuada de evaluación propuesta por la ecuación 5, además, como función de evaluación extra se asigna el nivel de retorno esperado dentro de un umbral específico móvil que asegura la generación completa de la frontera de eficiencia, es por tanto una función de evaluación multiobjetivo.

4.3. Selección

De la población completa se selecciona a los individuos que hayan obtenido mejor puntuación en la función de evaluación y algunos individuos al azar que a su vez pasaran a un registro histórico llamado salón de la fama, para este caso se usa un sistema de selección basado en “rueda de la fortuna.”

4.4. Recombinación

Dos puntos de cruce son seleccionados aleatoriamente (Fig. 2) de cada pareja de padres, los cuales intercambian las cadenas de genes entre ellos y producen la siguiente

generación, la probabilidad de cada pareja de padres de tener crías es comúnmente proporcional al desempeño de cada padre y sustituyen a los individuos peor evaluados de la siguiente generación.

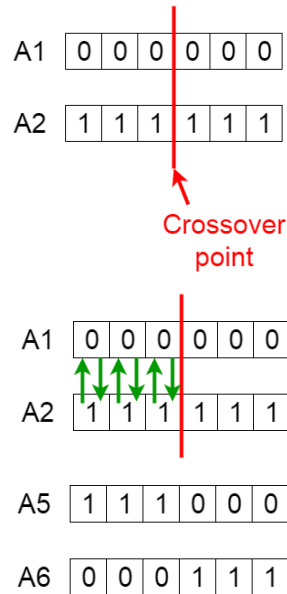


Fig. 2. Recombinación de genes sencilla.

4.5. Mutación

Cada integrante de la nueva generación tiene una probabilidad de que sus genes sean alterados con una probabilidad muy baja, para este caso de .3 con una función gaussiana con promedio en el mismo nivel del valor del gen específico y $\sigma=1$ (Fig. 3).

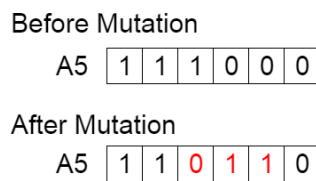


Fig. 3. Mutación sencilla para genes binarios.

4.6. Terminación

Es necesario crear un criterio de finalización, el cual, hasta no ser cumplido, implicará la repetición de todos los pasos anteriores, lo que creará una generación por cada iteración hasta obtener resultados satisfactorios, el criterio del presente algoritmo es la repetición hasta no mejorar el algoritmo en 10 iteraciones consecutivas. En la Fig. 4 se presenta el esquema general del algoritmo genético:

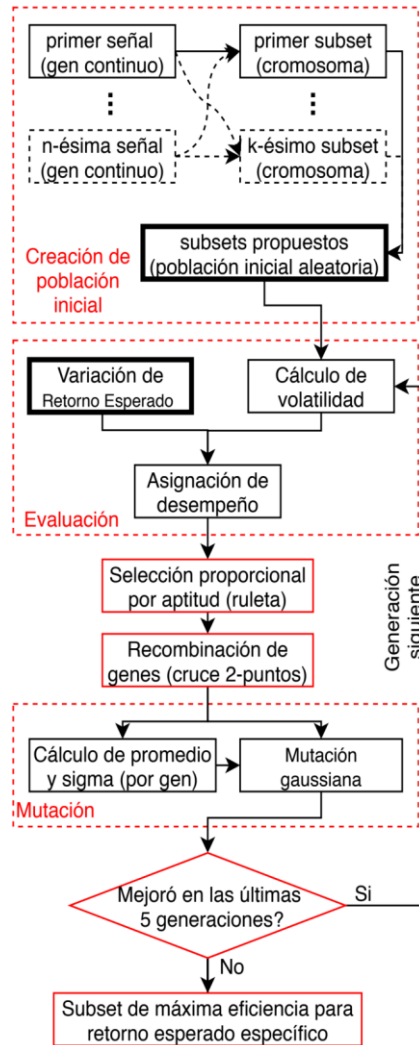


Fig. 4. Diagrama general de algoritmo genético.

5. Resultados

La frontera de eficiencia se dibuja a sí misma al intentar obtener las combinaciones de señales donde la volatilidad sea mínima en distintos niveles de retorno esperado, este borde contiene a los mejores conjuntos de señales e inutiliza todo lo que se aleje de él. No sería apropiado elegir combinaciones con el mismo nivel de retorno esperado pero con mayor riesgo o el mismo nivel de riesgo pero con menor retorno esperado.

Dependiendo del nivel de aversión al riesgo, cualquier punto (combinación obtenida) ubicado en la frontera de eficiencia que se encuentre entre los triángulos azul (mínimo riesgo) y rojo (máximo retorno) puede representar una opción válida de

selección de señales dentro del universo de configuraciones posibles.

En la Fig. 5 se observan los resultados de ejecutar el algoritmo para 10 señales de stocks incluidos en S&P 500.

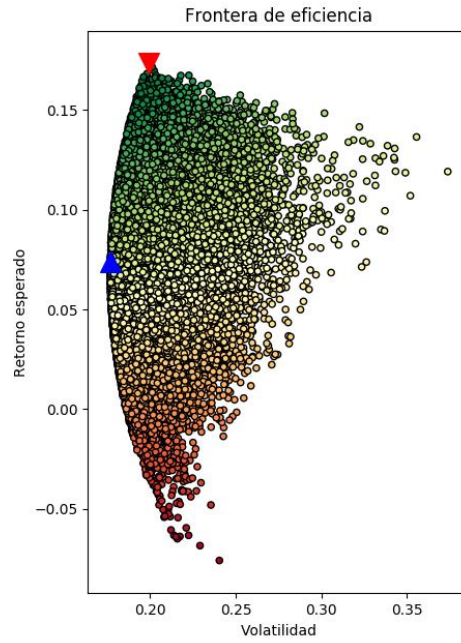


Fig. 5. Frontera de eficiencia obtenida para S&P 500.

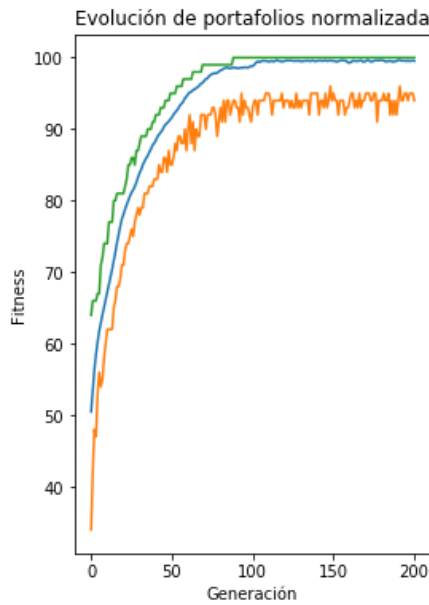


Fig. 6. Evolución normalizada de portafolio para riesgo determinado.

6. Conclusiones

Se propuso un método para explorar las combinaciones posibles de un conjunto de señales financieras con resultados satisfactorios, se observa una aproximación correcta a la frontera de eficiencia por lo que se cumple el objetivo.

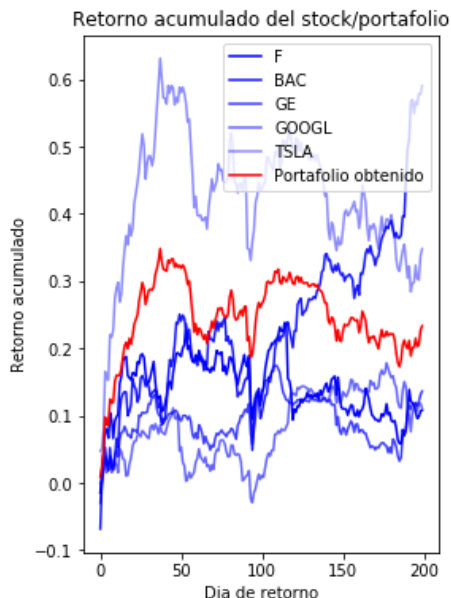


Fig. 7. Retorno de inversión de distintos stocks y portafolio propuesto.

Referencias

1. Markowitz, H.: Portfolio selection. *The Journal of Finance*, vol. 7, pp. 77–91 (1952)
2. Sharpe, W.F.: The sharpe ratio. *The journal of portfolio management*, vol. 21, pp. 49–58 (1994)
3. Li, B., Hoi, C.H.: Online portfolio selection: A survey. *ACM Computing Surveys*, vol. 46 (2014)
4. Potter, M.A., De Jong, K.A.: Cooperative Coevolution: An Architecture for Evolving Co-adapted Subcomponents. *Evolutionary Computation* 8(1), 1–29 (2000)
5. Pereira, R.: Genetic Algorithm Optimization for Finance and Investments. MPRA Paper University Library of Munich, pp. 1–26 (2000)
6. Fama, E.F.: Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics*, vol. 49, pp. 283–306 (1999)
7. Chen, Y., Yang, B., Abraham, A.: Flexible neural trees ensemble for stock index modeling. *Neurocomputing*, vol. 70, pp. 4–6 (2007)
8. Demiguel, V., Garlappi, L., Uppal, R.: How Inefficient is the 1/N Asset-Allocation Strategy? (2005)
9. Deb, K., Agrawal, S., Pratap, A., Meyarivan, T.: A Fast and Elitist Multi-Objective Genetic Algorithm: NSGA-II. *IEEE Trans. Evolut. Comput.* 6(2), 182–197 (2002)