Práctica 6: Metaheurísticas poblacionales

Víctor Tirado

victortiradoarregui@uma.es
Algoritmos de Búsqueda y Optimización Computacional. Universidad de Málaga.

En esta práctica vamos a aplicar técnicas avanzadas de optimización para resolver problemas reales bajo restricciones complejas. En particular, se trabajará con datos financieros reales para construir carteras de inversión eficientes, enfrentando el desafío de seleccionar un subconjunto limitado de activos mediante algoritmos evolutivos.

1 Introducción teórica

El problema de selección óptima de carteras es fundamental en finanzas cuantitativas. El modelo cásico de Markowitz propone una formulación basada en el equilibrio entre rentabilidad esperada y riesgo, donde el riesgo se cuantifica mediante la varianza de los retornos. En este contexto, una cartera **eficiente** es aquella que, para un nivel de riesgo dado, maximiza la rentabilidad, o viceversa.

En esta práctica nos enfrentamos a una versión restringida del modelo de Markowitz, que incorpora dos tipos de restricciones:

- **Restricción de presupuesto**: los pesos de los activos deben sumar exactamente uno ($\sum w_i = 1$).
- **Restricción de cardinalidad**: solo un número máximo K de activos pueden tener un peso positivo.

Estas restricciones reflejan condiciones reales de inversión, como la necesidad de diversificación controlada y la imposibilidad de fraccionar demasiado la cartera por motivos operativos o regulatorios. La formulación completa del problema es:

$$\min_{w \in R^S} \quad \frac{1}{2} w^T \Sigma w - \lambda w^T \mu, \quad sujetoa \quad \sum w_i = 1,\ \|w\|_0 \leq K$$

Resolver este problema no es trivial. Mientras que la versión sin restricciones se puede resolver con programación cuadrática convexa, la presencia de la norma ℓ_0 convierte el problema en **NP-difícil**. Para afrontar esta complejidad, se recurre a técnicas metaheurísticas como **CMA-ES (Covariance Matrix Adaptation Evolution Strategy)**, un algoritmo evolutivo robusto para optimización continua sin derivadas.

2 Descripción y preparación de los datos

Para la experimentación henos seleccionado 20 activos representativos del S&P100, como por ejemplo AAPL, GOOGL, TSLA, IBM, KO, etc. Descargamos sus precios diarios para el periodo de **entrenamiento** (1 de enero al 31 de diciembre de 2024) y para el periodo de **validación** (1 de enero al 28 de febrero de 2025) utilizando la librería `yfinance`.

Procesamos los datos de la siguiente manera:

1. Eliminamos columnas con datos faltantes (NaN).
2. Calculamos rentabilidades diarias simples:
$$R_{t,i} = \frac{P_{t,i} - P_{t-1,i}}{P_{t-1,i}}$$
3. Calculamos:
   - μ: vector de medias de rentabilidad para cada activo.
   - Σ: matriz de covarianzas de los retornos diarios.

3 Algoritmo CMA-ES modificado

3.1 Motivación del uso de CMA-ES

CMA-ES es especialmente adecuado para problemas de optimización en espacios continuos de alta dimensionalidad, donde la derivada de la función objetivo no está disponible o es costosa de calcular.

3.2 Función de proyección y reparación

Para manejar la restricción de cardinalidad hemos implementado un método de proyección que:

- Eliminanos pesos negativos.
- Conservamos los K mayores pesos.
- Normalizamos para cumplir la restricción.

3.3 Función objetivo

La función objetivo combina riesgo y rentabilidad tras aplicar la proyección:

$$f(w) = \frac{1}{2} w^T \Sigma w - \lambda w^T \mu$$

Donde w ha sido previamente proyectado para respetar la cardinalidad. Hemos evaluado distintas funciones de objetivos en el que incluimos penalizaciones pero no encontramos una mejora significativa.

3.4 Configuración del algoritmo

- Población inicial: vector uniforme de pesos.
- Dispersión inicial: $\sigma = 0.2$.
- Tamañoo población: 20.
- Criterios de parada: 100 iteraciones o convergencia.

4 Resultados experimentales

4.1 Pesos de la cartera óptima

Ejemplo con $K = 8$ y $\lambda = 10$:

```
AAPL: 0.0000  AMZN: 0.0000  CSCO: 0.0000  ...
NVDA: 1.0000  Resto: 0.0000
```

El resultado muestra un claro sobreajuste a un solo activo, en este caso NVIDIA. Esto indica que la proyección dura empobrece la diversidad del portafolio.

4.2 Evaluación en el periodo de validación

- Rentabilidad media diaria: -0.0028
- Volatilidad (desviación estándar): 0.0452
- Ratio de Sharpe: -0.0610

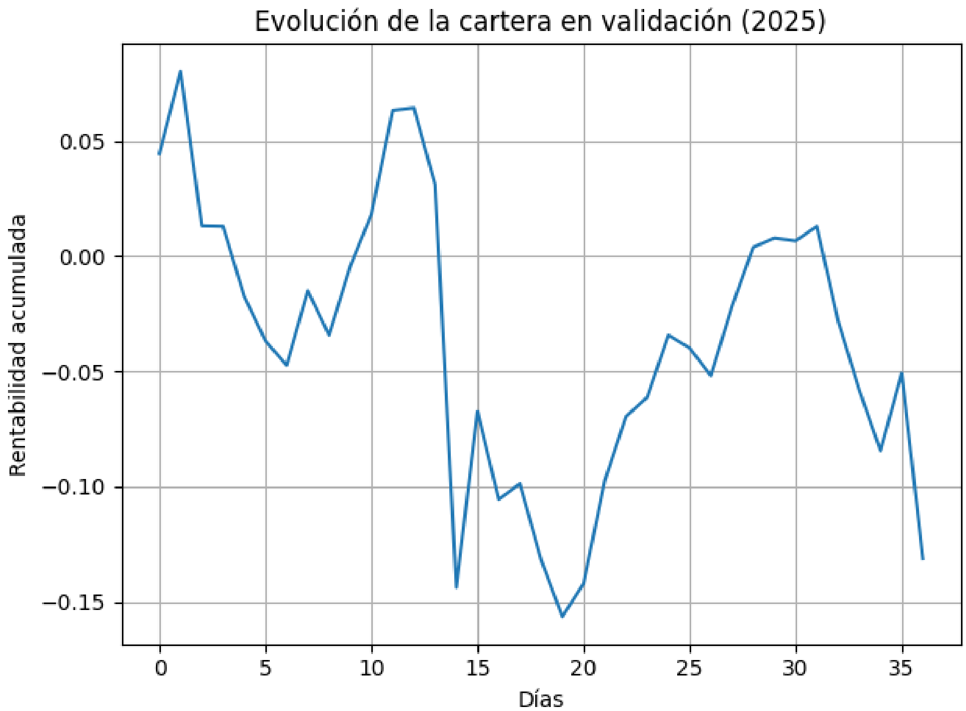

Fig. 1. Evolución de la rentabilidad acumulada de la cartera óptima durante enero-febrero de 2025.

5 Conclusiones sobre los diferentes parámetros del modelo y valores de $\boldsymbol{\lambda}$

Durante las pruebas experimentales, hemos evaluado distintos valores del parámetro λ (por ejemplo, $\lambda = 1$, $\lambda = 10$, $\lambda = 50$), observando los siguientes comportamientos:

- Con valores bajos de λ (e.g., $\lambda = 1$), el modelo prioriza fuertemente la rentabilidad, lo que puede llevar a soluciones arriesgadas y menos estables.
- A medida que λ aumenta, se incrementa el peso relativo de la varianza, y por tanto el modelo tiende a seleccionar activos más estables aunque menos rentables.
- En valores altos como $\lambda = 50$, la cartera final suele estar dominada por activos de baja varianza, a costa de una posible caída en rentabilidad.
- Se observó que $\lambda = 10$ ofrece un compromiso razonable entre riesgo y retorno, aunque el resultado final sigue altamente influido por la estructura de la matriz de covarianzas.

Estos experimentos muestran que λ actúa como un control directo del riesgo del inversor, y su elección debe hacerse en función del perfil deseado para la cartera final.

6 Mejora del modelo

Problema observado: el modelo converge rápidamente a soluciones triviales con un único activo, en este caso siempre converge a NVIDIA. Esto sucede porque la proyección fuerza un espacio factible extremadamente reducido, y el algoritmo converge sin explorar nuevas combinaciones.

Alternativas propuestas:

- Sustituir proyección por penalización suave proporcional al número de activos activos.

- Forzar diversidad con restricciones adicionales.
- Usar estrategias mixtas.
- Introducir un operador de mutación personalizado para explorar combinaciones más amplias.

7 Conclusiones

- La proyección dura es sencilla de implementar pero puede inducir soluciones no diversificadas.
- CMA-ES es eficaz optimizando en espacios restringidos, pero necesita ajustes para evitar soluciones degeneradas.
- La validación empírica muestra que sobreajustar a un activo puede llevar a bajo rendimiento fuera de muestra.
- El equilibrio entre interpretabilidad, diversificación y rendimiento es clave en modelos financieros basados en metaheurísticas.