Un modelo de selección de carteras basado en la predicción de precios

E Vercher, A Corberán, JV Segura, JD Bermúdez

Departamento de Estadística e I.O. Universitat de València
Centro de Investigación Operativa. Universidad Miguel Hernández de Elche

XIII Reunión Grupo Español de Decisión Multicriterio
Donostia - San Sebastián, 23 de julio de 2021

Esquema

1. *El problema de selección de carteras*
   - Datos
   - Introducción
2. *Valor de una cartera*
   - Modelización y predicción
3. *Modelos y algoritmos*
   - Resultados

Precios de activos de la Bolsa de NY

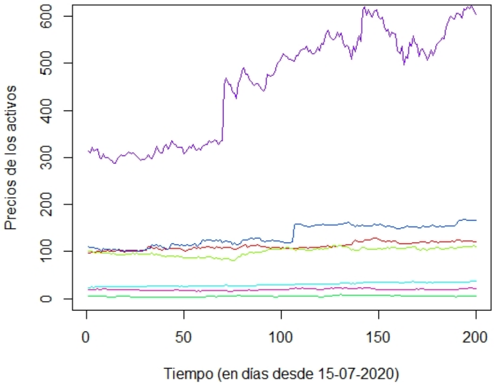
Objetivo: seleccionar carteras eficientes que equilibren el rendimiento-riesgo, y que atiendan los requisitos del inversor.

- Se consideran los precios diarios al cierre de $n = 72$ activos que cotizan en la Bolsa de Nueva York (días laborables).
- Los activos pertenecen a los sectores: Health Care Equipment & Services y Pharmaceuticals & Biotechnology.
- El precio de un activo $i = 1, ..., n$ en el periodo $t = 1, ..., T$ se define como y_{it}, desde el 15 de agosto de 2020 (t=1) al 28 de abril de 2021 (T=200).
- El rendimiento del activo i-ésimo en el periodo t se calcula como:

$$r_{it} = \frac{y_{i(t+1)} - y_{it}}{y_{it}} \times 100$$

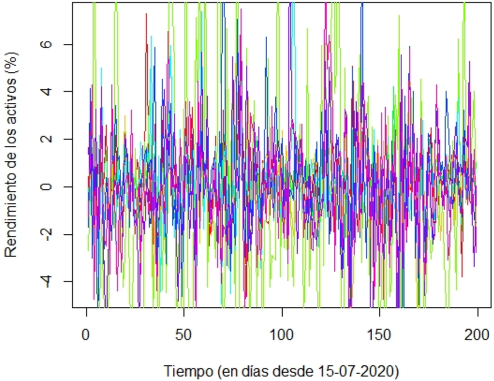
Gráfica temporal de precios

Ejemplo: Se han ordenados alfabéticamente los 72 activos y hemos seleccionado los 8 primeros. La gráfica muestra la serie de precios diarios al cierre.



Gráfica temporal de los rendimentos

Ejemplo: La gráfica de los rendimientos de los 8 activos seleccionados muestra que algunos de ellos presentan una mayor volatilidad.



Selección de carteras

- Una cartera es una colección de activos financieros en los que se invierte un capital.
- El problema de seleccionar una cartera consiste en elegir los activos y las proporciones que proporcionen una inversión óptima para un inversor dado.
- Histórico de datos: precios de los activos
  - Se asume que los rendimientos de cada activo son observaciones independientes, cuya relación se resume en la matriz de covarianzas.
  - Se asume que la distribución conjunta de los rendimientos es Normal multivariante.
- Modelo de media-varianza (MV: Markowitz,1952)
  - Maximizar el rendimiento esperado fijando el riesgo máximo que el inversor está dispuesto a asumir, y/o
  - Minimizar el riesgo de la inversión fijando el minimo rendimiento esperado que el inversor desea alcanzar.

Decisión multicriterio

- Alternativas al modelo clásico de MV
  - Introducir momentos de orden superior como objetivos y/o restricciones.
  - Utilizar lógica fuzzy para aproximar la incertidumbre del rendimiento de los activos, definiendo nuevos modelos de selección.
  - Aproximar la incertidumbre del rendimiento futuro utilizando el rendimiento histórico de cada cartera, permitiendo incorporar información implícita de las relaciones contemporáneas de los activos.
- Respecto de la independencia entre observaciones
  - Considerar las series temporales de precios de los activos individualmente para seleccionar aquellos de los que se espera un mejor comportamiento, y construir una cartera con ellos.
  - Considerar el valor de cada cartera como una serie temporal, y utilizar las predicciones de su valor para estimar los parámetros de interés sobre el riesgo y rendimiento de la misma.

Para todas estas modelizaciones del problema de selección de carteras basados en el equilibro riesgo-rendimiento, la optimización multi-objetivo permite determinar la frontera de Pareto y seleccionar la cartera que satisfaga los requisitos del inversor.

Serie temporal del valor de una cartera

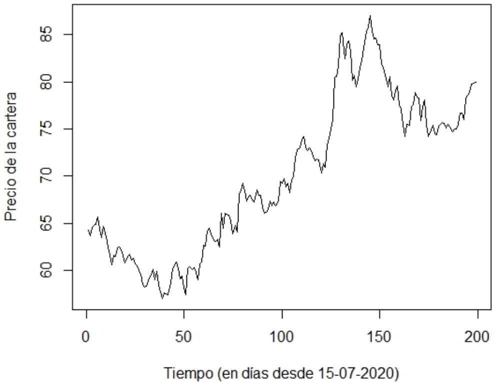
- Consideremos un mercado con n activos que han cotizado en $t = 1, .., T$. Para cada activo i, $i = 1, 2, ...n$, consideramos la serie temporal de sus precios en el mercado $\{y_{i1}, y_{i2}, ...y_{iT}\}$.
- Una cartera en el instante T consiste en un vector de pesos no negativos $X = (x_1, x_2, ..., x_n)$, siendo x_i la proporción del capital M invertido en el activo i, en tantos por uno.
- Para una cartera X en T, la cantidad invertida en i es:

$$s_i = M \times \frac{x_i}{y_{iT}}$$

- Si se asume que las proporciones asignadas a cada activo se han mantenido constantes a lo largo de todo el periodo observado, el valor de la cartera puede considerarse como la serie temporal $\mathbf{V(X)} = \{v_1(X), v_2(X), ..., v_T(X)\}$, siendo $v_t(X) = y_{1t} \times s_1 + y_{2t} \times s_2 + ... + y_{nt} \times s_n$, para $t = 1, 2, ..., T$.

Serie temporal de una cartera

- La gráfica muestra la serie temporal del valor de una cartera naïf construida con una inversión de 10 euros en cada uno de los 8 activos del ejemplo, en $T = 200$.
- El valor de la cartera $V(X)$ se calcula hacia atrás utilizando los vectores de precios.



Modelización del valor de una cartera

- La serie temporal **V(X)** $= \{v_t\}_{t=1}^{T}$ no presenta un patrón estacional.
- El futuro valor de la cartera X se predice mediante el modelo aditivo de Gardner con tendencia amortiguada. Los datos observados se describen como:

$$v_t = a_{t-1} + \phi b_{t-1} + \epsilon_t$$

  siendo $0 < \phi < 1$ el parámetro de amortiguación, a_t y b_t el nivel y la tendencia y ϵ_t el error (iid Normal(0,σ^2)) en t.
- Las ecuaciones de transición son:

$$\begin{array}{rcccl} a_t & = & \alpha v_t + (1-\alpha)(a_{t-1} + \phi b_{t-1}) & = & a_{t-1} + \phi b_{t-1} + \alpha \epsilon_t \\ b_t & = & \beta(a_t - a_{t-1}) + (1-\beta)\phi b_{t-1} & = & \phi b_{t-1} + \alpha\beta\epsilon_t \end{array}$$

  siendo $0 < \alpha, \beta < 1$ los parámetros de suavizado.
- El estimador maximoverosímil de $\theta = (\alpha, \beta, \phi)$ y de (a_0, b_0) se obtiene resolviendo un problema de optimización no lineal.

Predicción del valor de una cartera

- Para la serie $\mathbf{V(X)} = \{v_t\}_{t=1}^{T}$, el modelo anterior de suavizado exponencial puede ser formulado matricialmente como:

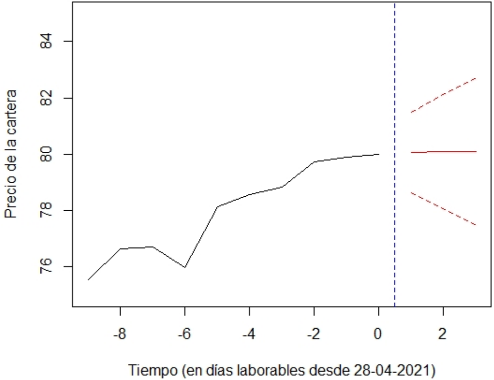
$$\mathbf{V(X)} = A\,\omega + L\,\varepsilon$$

- La distribución conjunta del vector **V(X)** es Normal multivariante, $E(\mathbf{V(X)}) = A\,\omega$ y $\sum_{\mathbf{V(X)}} = \sigma^2 LL'$. Los estimadores de ω y σ^2 se obtienen mediante expresiones analíticas a partir de la estimación de θ, $\hat{\theta}$.
- Sea $\mathbf{V}_f(\mathbf{X}) = \{v_t\}_{t=T+1}^{T+h}$ el vector de futuros valores de la cartera X, y asumimos que el vector $(\mathbf{V}, \mathbf{V}_f)(\mathbf{X})$ sigue el mismo modelo lineal.
- La estimación puntual del valor futuro de la cartera sería la esperanza $E(\mathbf{V}_f|\mathbf{V})$, para un valor dado de $t = T+1, .., T+h$.
- El rendimiento esperado de la cartera X a h-etapas sería:

$$E(r_h(X)) = \frac{E(v_{T+h}) - v_T}{v_T}$$

Predicción del precio de la cartera

- Para la cartera naïf generada con 8 activos, se presentan las predicciones de su valor para los tres días siguientes, con intervalos al 80 %.
- El último día observado (28-04-2021) corresponde al 0 en la gráfica.



Objetivos y restricciones

- Funciones objetivo: $z_i = f_i(\mathbf{X}, \mathbf{V})$, $i = 1, 2, .., m$
  - Valor esperado de la cartera: $z_1 = E(v_{T+k}(X))$, $k = 1, ..., h$.
  - Variabilidad del valor de la cartera: $z_2 = Var(v_{T+k}(X))$, $k = 1, ..., h$.
  - Probabilidad de obtener un rendimiento negativo: $Pr(r_k(X) \leq 0)$, $k = 1, ..., h$.
- Restricciones: $X \in \mathfrak{F}$
  - $x_1 + \cdots + x_n = 1$, con $x_i \geq 0 \ \ \forall i$.
  - $l_i \leq x_i \leq u_i$, $i = 1, \ldots, n$
  - Cardinalidad restringida: $K_l \leq \sum(x_i > 0) \leq K_u$.
  - Tendencia positiva en T: $b_T(X) > 0$.
  - Percentiles positivos del intervalo de predicción del valor de la cartera en $k = 1, ..., h$.
  - Cota inferior para el Sharpe ratio de X, en $T + 1, ..., T + h$.

Modelo MV para el valor de la cartera

Para generar la frontera de Pareto con carteras eficientes para los objetivos valor de la cartera-riesgo, planteamos un problema de optimización bi-objetivo:

- $(MV(v_{T+1}))$: Modelo *Media-Varianza del valor en* $T+1$.

$$
\begin{array}{lll}
\mathit{Max} & E(v_{T+1}(X)) & \text{(Valor esperado)} \\
\mathit{Min} & \sigma(v_{T+1}(X)) & \text{(Riesgo)} \\
\mathit{s.a.} & x_1 + \cdots + x_n = 1 & \\
 & 0 \le x_i \le 0{,}3 & i = 1, \ldots, n \\
 & 6 \le \sum(x_i > 0) \le 9 &
\end{array}
$$

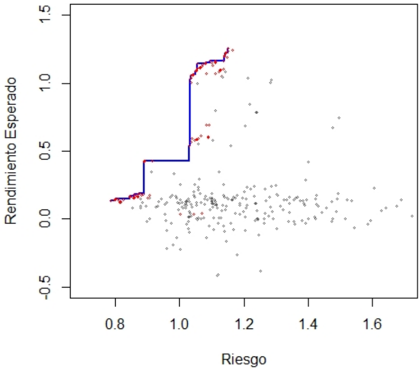
- Resolvemos el problema (NP-duro) aplicando un algoritmo genético, que hemos modificado para mejorar la construcción del frente de Pareto, y para disminuir los tiempos de computación.

Algoritmo genético

- Nuestro algoritmo genético (GA), genera una población inicial utilizando la distribución de Dirichlet en el simplex de orden K, siendo $K = \{6, 7, 8, 9\}$ el tamaño de la cartera.
- En cada generación el algoritmo:
  - Encuentra las soluciones no dominadas para construir la frontera superior.
  - Selecciona las carteras élite (20 %) y se les aplica aleatoriamente un operador mutación que perturba un par de proporciones. Comprueba la cardinalidad de la solución.
  - Se ha introducido un operador de preservación de élites que incorpora la información de preferencias del inversor.
  - Criterio de parada: 10 generaciones o distancia entre fronteras superiores sucesivas menor que 10^{-5}.
- Se generan diversas poblaciones en paralelo con semillas diferentes, a las que se les aplica el GA.
- Se construye una última población mediante la unión de las que se obtienen como población final en los lanzamientos en paralelo. Se aplica de nuevo el GA.

Aproximación de la frontera de Pareto

- Para resolver el problema (MV(v_{T+1})), se ha lanzado el algoritmo genético tres veces en paralelo, para una población de 40 individuos.
- En gris las poblaciones iniciales. En rojo las poblaciones finales de los tres lanzamientos. La frontera de Pareto aproximada en azul.



Conclusiones

- Hemos definido la serie temporal del valor de una cartera $V(X)$ en un instante T, utilizando retrospectivamente el histórico de los precios de los activos que la componen.
- Hemos aplicado un modelo de suavizado exponencial para predecir los futuros valores de la cartera.
- Para obtener carteras eficientes, hemos propuesto un modelo de optimización bi-objetivo basado en las predicciones del valor/precio de las carteras.
- Hemos modificado un GA para que las carteras eficientes satisfagan los requisitos impuestos por el inversor.

Esta es una primera aproximación al problema planteado. Quedan muchas cuestiones por abordar y problemas por resolver relacionados con esta propuesta.