Proyecto Estadística y Optimización

Máster en Ciencia de Datos - UV

Adrián Carrasco Alcalá, Javier Herrero Pérez y Clara Montalvá Barcenilla

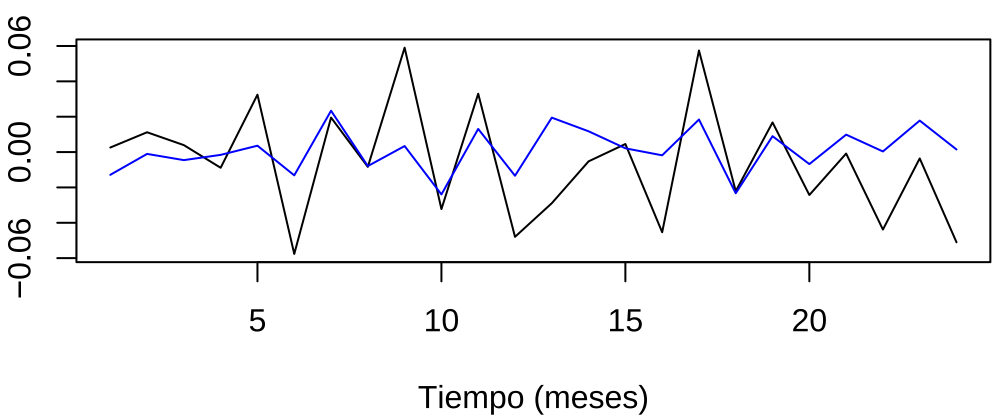
Curso 2025-2026

Organización del trabajo

El trabajo se ha dividido entre los miembros del grupo de la siguiente manera:

- Javier Herrero Pérez: Apartado 2 y 3.1
- Clara Montalvá Barcenilla: Apartado 3.2
- Adrián Carrasco Alcalá: Apartado 4

Apartado 2



La optimización de la cartera en cada periodo t requiere estimaciones del rendimiento esperado μ_t y de la matriz de covarianza Σ_t para los $N = 5$ activos. Para ello, se ha utilizado un enfoque de modelos de series temporales aplicando modelos autoregresivos ARIMA. Más concretamente, para cada uno de los cinco activos se ha adoptado un modelo ARIMA y se han ido estudiando los valores a partir de una función oneStepAhead.

Los datos de los que disponemos se extienden a lo largo de 10 años (120 periodos), por lo que se ha decidido entrenar con 8 años y los otros 2 años como valores de test. La figura anterior permite evaluar la calidad del modelo de predicción ARIMA en el periodo de prueba. La gráfica muestra la comparación de los rendimientos reales frente a los rendimientos pronosticados para un solo acivo durante el periodo de prueba. Se observa que aunque la amplitud no sea perfecta, sí que es capaz de captar correctamente cuando sube y cuando baja.

El resultado del RMSE promedio es del 0.0175 (1.75%) en el pronóstico de los 5 activos. Dado el comportamiento ruidoso de las series este resultado indica que el modelo ARIMA ha logrado capturar la componente predecible de la señal. La magnitud del error se debe principalmente a la componente impredecible.

Apartado 3.1

La función de utilidad media-varianza es $U_{MV}(\alpha_t, \gamma; \mu_t, \Sigma_t) = \alpha_t^T \mu_t - \frac{\gamma}{2}\alpha_t^T \sum_t \alpha$.

Esta función se maximiza, esto equivale a minimizar $-U_{MV}$ sujeto a la restricción $\sum \alpha_t = 1$. Se trata de un problema de programación cuadrática sin restricciones de desigualdad, entonces el problema a resolver es el siguiente:

$$\min_{\alpha}\{\frac{\gamma}{2}\alpha_t^T \sum_t \alpha_t - \alpha_t^T \mu_t\} \quad \text{s.a} \quad \sum \alpha_t = 1$$

En este caso, donde se permiten las posiciones cortas, la solución óptima α^* se puede obtener de manera teórica mediante el uso de multiplicadores de Lagrange. Se define el Lagrangiano L de la siguiente forma:

$$L = \frac{\gamma}{2}\alpha_t^T \sum_t \alpha_t - \alpha_t^T \mu_t + \lambda(\mathbf{1}^T \cdot \alpha - 1)$$

donde $\mathbf{1}^T$ es un vector columna de 1 con N filas, donde N es el número de activos que en este problema es 5. Se resuelve el sistema:

$$\frac{\partial L}{\partial \alpha} = \gamma \Sigma \alpha^* - \mu + \lambda^* \mathbf{1} = 0$$
$$\frac{\partial L}{\partial \lambda} = \mathbf{1}^T \alpha^* - 1 = 0$$

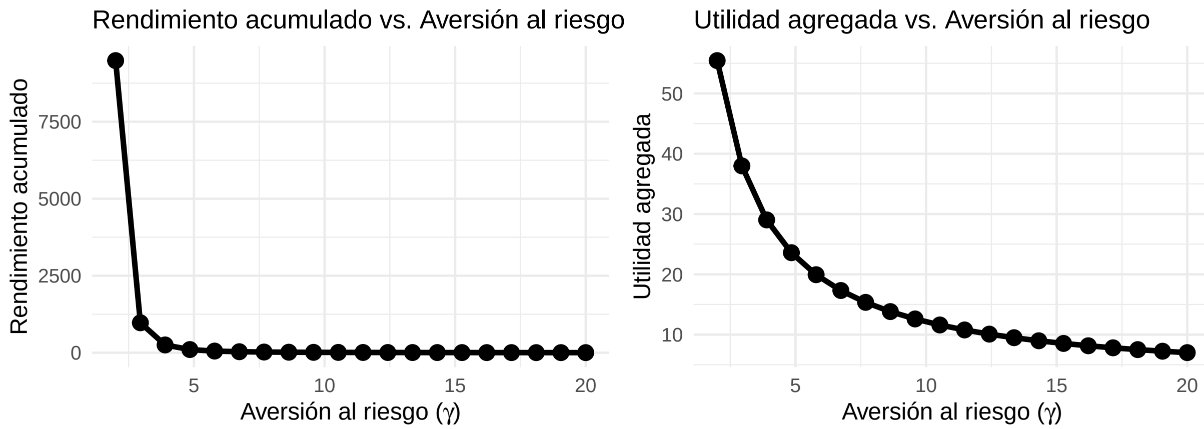
donde α^*, λ^* las soluciones del sistema.

Sustituyendo se obtiene lo siguientes resultados:

$$\lambda^* = \frac{\mathbf{1}^T \Sigma^{-1} \mu - \gamma}{\mathbf{1}^T \Sigma^{-1} \mathbf{1}}$$
$$\alpha^* = \frac{1}{\gamma} \Sigma^{-1} (\mu - \lambda^* \mathbf{1})$$

Esta solución analítica se implementa directamente en la función `getAlphaMV`, lo que permite obtener el valor exacto de la asignación del portafolio α.

Para implementar la solución es necesario obtener la matriz de covarianza Σ_t que se ha calculado previamente en `getSigmaMV`, este enfoque justifica el uso de los modelos ARIMA para la volatilidd (elementos diagonales de Σ_t) mientra que la dependencia entre los activos se saca de la historia de los datos.



La gráfica muestra la relación entre el retorno acumulado y el parámetro de riesgo γ. El comportamiento genera que se observa es que el retorno cae drásticamente a medida que γ aumenta hasta que se estabiliza. Tomar un valor pequeño de γ conlleva tomar mucho riesgo, provocando un mayor retorno acumulado.

- Riesgo mínimo: Elegir valores altos como $\gamma = 20$ significa priorizar la minimización extrema del riesgo. Aunque esto resulta en cartera más segura, se observa que se penaliza excesivamente el rendimiento, logrando un retorno acumulado de 295% y una utilidad promedio de $U_{MV} = 7.011$.
- Balance óptimo: Si se busca un balance eficiente entre riesgo y retorno el valor de $\gamma = 5$ es un punto destacable en la curva de la gráfica. Este valor impone un riesgo considerable y un retorno grande 9107% lo que resulta en una utilidad agregada mucho mayor $U_{MV} = 22.88$.

Elegir valores $\gamma < 5$ conllevaría un riesgo mucho mayor y a su vez un rendimiento elevado, si se busca un equilibrio entre rendimiento y seguridad el valor de $\gamma = 5$ sería adecuado, pero si se busca maximizar el retorno el valor de $\gamma = 1$ sería el adecuado.

Como conclusión se elige $\gamma = 5$ como el valor de riesgo ya que busca una utilidad agregada equilibrada como se observa en la figura de la derecha, manteniendo un rendimiento alto mientras que mantiene el riesgo bajo control según la curva mostrada.

Apartado 3.2

El objetivo sigue siendo maximizar la función de utilidad Media-Varianza, pero a diferencia del apartado 3.1, en este caso se prohíben las posiciones cortas, lo que introduce una restricción de desigualdad:

$$\alpha_{t,i} \geq 0 \quad \forall\, i = 1, ..., 5$$

Al combinar esta restricción con la restricción de presupuesto (suma de ponderaciones igual a 1), el problema de optimización cuadrática a resolver es el siguiente:

$$\min_{\alpha} \left\{ \frac{\gamma}{2} \alpha_t^T \Sigma_t \alpha_t - \alpha_t^T \mu_t \right\} \quad \text{s.a} \quad \sum_i \alpha_{t,i} = 1 \quad \text{y} \quad \alpha_{t,i} \geq 0$$

Debido a la presencia de la restricción de no-negatividad ($\alpha_i \geq 0$), la solución analítica con multiplicadores de Lagrange (Kuhn-Tucker) es más compleja y generalmente se aborda mediante métodos numéricos. Para resolver este problema de programación cuadrática de manera eficiente, se ha utilizado la función `solve.QP` de la librería de `quadprog`. La función `getAlphaMVPos` implementa los siguientes componentes requeridos por `quadprog`:

- Matriz D: matriz $D = \gamma\Sigma_t$, que define la parte cuadrática del problema.
- Vector d: vector $d = \mu_t$, que define la parte lineal.
- Matriz de Restricciones A: contiene las restricciones de igualdad y desigualdad.

1. Igualdad: $\sum_i \alpha_{t,i} = 1$.
2. Desigualdad: $\alpha_{t,i} \geq 0$, representada por la matriz identidad $\mathbf{I}$ (A_ineq <- diag(n)).

- Vector b: contiene los valores de las restricciones (tanto de igualdad como de desigualdad).

La solución obtenida proporciona la asignación óptima α^* para cada periodo t sujeta a la prohibición de posiciones cortas.

La estimación de la matriz de covarianza Σ_t para este apartado (`getSigmaMVPos`) se realiza con el mismo enfoque que en la Sección 3.1. La volatilidad individual (elementos diagonales $\sigma_{i,t}$) se obtiene de los modelos ARIMA del Apartado 2. La matriz de correlación (R) se estima a partir del historial de rendimientos y la matriz de covarianza se construye mediante la descomposición $\Sigma_t = DRD$.

Apartado 4

En este caso, la función de utilidad que se utiliza es la función de log-utilidad ($U_{\log}$), la cual reconoce la naturaleza compuesta del rendimiento de un portafolio. Su expresión durante un periodo es:

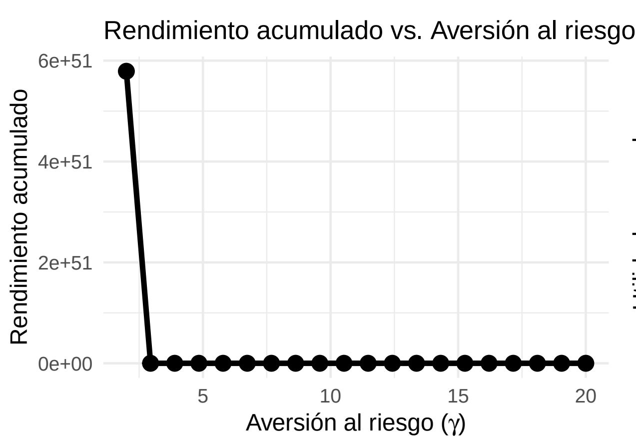
$$U_{\log}(\alpha_t, \gamma; \mu_t, \Sigma_t) = \log(1 + \alpha^T \mu_t) - \frac{\gamma}{2} \frac{\alpha_t^T \Sigma_t \alpha_t}{(1 + \alpha_t^T \mu_t)^2}.$$

Se busca maximizar dicha función, lo cual equivale al problema de minimizar $-U_{\log}$ sujeto a la restricción $\sum \alpha_t = 1$. Por tanto, el problema a resolver es el siguiente:

$$\min_{\alpha} \left\{ \log(1 + \alpha^T \mu_t) - \frac{\gamma}{2} \frac{\alpha_t^T \Sigma_t \alpha_t}{(1 + \alpha_t^T \mu_t)^2} \right\} \quad \text{s.a} \quad \sum \alpha_t = 1.$$

Al igual que ocurría en el apartado 3.2, la solución analítica de este problema con multiplicadores de Lagrange vuelve a ser compleja, por lo que se debe recurrir de nuevo a métodos numéricos.

Para resolverlo, se ha utilizado la función `optim` de la librería de `stats`, partiendo de una asignación de pesos uniforme (mismo α para los 5 activos) y utilizando el método BFGS.



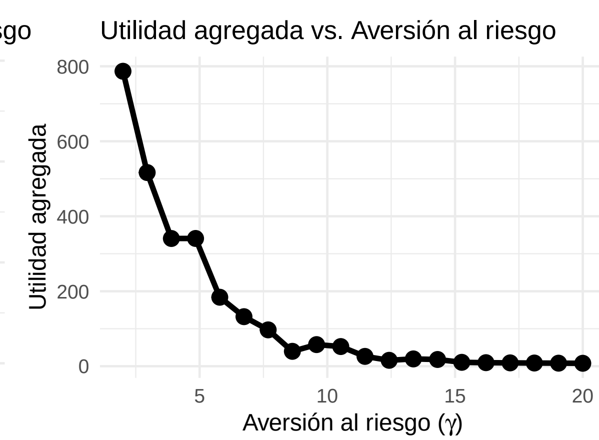


Table 1: Rendimiento y utilidad acumulada para cada gamma

Gamma	Rendimiento	Utilidad
2.00	5.7899e+51	786.78
2.95	-5.8231e+36	516.60
3.89	9.7493e+23	340.69
4.84	-1.7547e+24	340.80
5.79	8.4536e+13	183.98
6.74	1.5866e+11	132.15
7.68	9.9536e+08	96.96
8.63	8.2035e+03	39.72
9.58	2.5330e+04	57.66
10.53	1.1281e+04	52.47
11.47	4.0163e+02	26.23
12.42	6.8797e+01	15.77
13.37	1.3508e+02	19.49
14.32	1.1117e+02	17.93
15.26	2.3056e+01	10.40
16.21	9.0967e+00	9.45
17.16	7.2237e+00	8.94
18.11	6.0521e+00	8.51
19.05	5.2214e+00	8.14
20.00	4.5998e+00	7.82

El comportamiento que se observa es extremo y atípico en comparación con una curva teórica suave. Se obtienen valores extremadamente grandes que caen rápidamente a 0. La función de utilidad agregada, sin embargo, sí que presenta un comportamiento relativamente más suave.

Para valores γ inferiores a 8, se tiene una zona de inestabilidad en la que la utilidad agregada es extremadamente volátil. Se obtiene, por tanto, un portafolio muy sensible y posiblemente inestable.

A partir $\gamma = 8$, la función de utilidad deja de decrecer a tanta velocidad, aunque no empieza a estabilizarse hasta que se alcanzan valores de $\gamma \geq 15$.

Con todo esto, encontrar un valor óptimo de γ en este caso no es tan sencillo. Asumir demasiado riesgo puede llevar a grandes pérdidas. Lo que parece claro es que se debería elegir un $\gamma \geq 8$, e incluso $\gamma \geq 15$, si lo que se prioriza es la seguridad en la inversión.

Es por ello que se escoge un factor de aversión al riesgo $\gamma = 15$.

Resultados

Visualicemos los resultados obtenidos para cada uno de los apartados a 24 meses vista en la siguiente tabla.

Table 2: Resultados finales tras 24 meses

RMSE	Rendimiento apartado 3.1 (retMV)	Utilidad apartado 3.1 (Umv)	Rendimiento apartado 3.2 (retMVPos)	Utilidad apartado 3.2 (UmvPos)	Rendimiento apartado 4 (retLog)	Utilidad apartado 4 (ULog)
0.0175	91.0723	22.8814	0.6212	3.1182	18.9136	10.4375