Trabajo práctico N4 - Optimización

Martina Grünewald y Victoria Lynch

3 de Noviembre 2023

1 Introducción

Este informe de investigación aborda dos problemas clave en la optimización y métodos numéricos: la resolución de sistemas lineales mediante el algoritmo de gradiente descendente y la optimización de portafolios de inversión utilizando el modelo de Markowitz. Ambos casos representan desafíos prácticos en finanzas y ciencia de datos, donde se busca minimizar funciones de costo o riesgo bajo ciertas restricciones.

En el primer caso, se aplicó el algoritmo de gradiente descendente para resolver sistemas lineales sobredeterminados, explorando la regularización L2 y comparando resultados con la solución exacta mediante SVD. En el segundo caso, se abordó la optimización de portafolios de inversión, utilizando el modelo de Markowitz para encontrar combinaciones óptimas de activos. Se analizaron las soluciones bajo distintas restricciones, incluyendo la asignación total de inversión y niveles mínimos de riesgo aceptable. Este informe ofreció una visión detallada de los métodos numéricos aplicados en sistemas lineales y la optimización de portafolios, subrayando su relevancia en la toma de decisiones financieras y análisis de datos.

2 Métodos numéricos utilizados

2.1 Gradiente decendente

El algoritmo de Gradiente Descendente es un método iterativo utilizado para encontrar el mínimo de una función. En el contexto de cuadrados mínimos, se busca minimizar la función de costo $F(x)$ dada por:

$$F(x) = (Ax - b)^T (Ax - b)$$

El algoritmo inicia con una condición inicial x_0 y se actualiza iterativamente utilizando la fórmula:

$$x_{k+1} = x_k - s\nabla F(x_k)$$

donde s es el tamaño del paso (tasa de aprendizaje) y $\nabla F(x_k)$ es el gradiente de F evaluado en x_k. El objetivo es converger hacia el valor de x que minimiza $F(x)$.

2.2 Regulación L2

La regularización L2 se incorpora a la función de costo para evitar el sobreajuste y promover soluciones más estables. La función de costo regularizada $F_2(x)$ se define como:

$$F_2(x) = F(x) + \delta_2 \|x\|_2^2$$

donde δ_2 es un parámetro de regularización y $\|x\|_2$ es la norma-2 del vector x. Esta adición penaliza soluciones con normas grandes, proporcionando una solución más suave y generalizable.

2.3 Descomposicion de valores Singulares (SVD)

La Descomposición de Valores Singulares (SVD) es una técnica que proporciona una solución exacta al sistema sobredeterminado $Ax = b$. Se descompone la matriz A en tres matrices U, S, y V^T, donde U y V^T son matrices ortogonales, y S es una matriz diagonal con los valores singulares de A. La solución exacta se obtiene como:

$$x_{\text{exact}} = V \cdot S^{-1} \cdot U^T \cdot b$$

La SVD es computacionalmente costosa pero proporciona una solución precisa para comparar con las soluciones aproximadas obtenidas mediante métodos iterativos como el Gradiente Descendente.

2.4 Matriz de covarianza

La matriz de covarianza es una matriz en la que , dados los datos de un conjunto X se realiza la covarianza entre cada elemento del conjunto. Su formula de la matriz es:

$$\text{cov}(X) = \frac{1}{n-1} \cdot (X - \bar{X})^T \cdot (X - \bar{X})$$

donde X es la matriz de datos, $\bar{X}$ es el vector de medias de las columnas de X y n es el número total de observaciones. Por su definición, la diagonal de la matriz de covarianza es la varianza de los elementos del conjunto X (porque es la covarianza de dos elementos iguales, que equivale a la varianza del elemento).

2.5 Sequential Least Squares Quadratic Programming (SLSQP)

El método de SLSQP es una implementación de optimización no lineal que busca encontrar el mínimo de una función objetivo sujeta a restricciones, utilizando técnicas de optimización cuadrática y mínimos cuadrados secuenciales. Utiliza restricciones de igualdad:

$$\text{constraint\_sum}(x) = \sum_{i=1}^{n} x_i - 1$$

Esta restricción garantiza que la suma de las variables de decisión x_i (representando el porcentaje de inversión en cada activo) sea igual a 1, es decir, la asignación total de inversión.
Y después están las restricciones de desigualdad:

$$\text{constraint\_risk}(x, r_{\text{min}}, \text{promedio\_returns}) = \text{promedio\_returns} \cdot x - r_{\text{min}}$$

que impone que el riesgo del portafolio, medido por el producto punto entre los retornos promedio y el vector de variables de decisión x no sea menor que un valor mínimo r_{min}.
r_{min} siendo el rendimiento mínimo esperado que se quiere alcanzar con el mix escogido para la inversión.

3 Desarrollo experimental

3.1 Cuadrados mínimos mediante descenso por gradiente

En la configuración del experimento, se procedió a la generación de matrices aleatorias A y vectores b en un contexto donde $A \in \mathbb{R}^{n \times d}$, $x \in \mathbb{R}^d$, y $b \in \mathbb{R}^n$. Para garantizar que el sistema fuera sobredeterminado, se fijaron $n = 5$ y $d = 100$. La función de costo $F(x)$ se definió de acuerdo con la ecuación:

$$F(x) = (Ax - b)^T (Ax - b)$$

expresando la meta de minimizar la discrepancia entre Ax y b.

La implementación del algoritmo de gradiente descendente siguió la fórmula iterativa:

$$x_{k+1} = x_k - s\nabla F(x_k)$$

utilizando $s = 1/\lambda_{\text{max}}$, donde λ_{max} representa el autovalor máximo de la matriz Hessiana H de F. La regularización L2 fue incorporada a la función de costo, dando origen a $F2(x)$ mediante la ecuación:

$$F2(x) = F(x) + \delta_2 \|x\|_2^2$$

con $\delta_2 = 10^{-2}\sigma_{\text{max}}$, siendo σ_{max} el valor singular máximo de A.

Durante el experimento, se varió el parámetro de regularización δ_2, y se realizaron comparaciones entre las soluciones obtenidas mediante el algoritmo de gradiente descendente y la Descomposición de Valores Singulares (SVD).

Estas comparaciones permitieron evaluar la sensibilidad del método a los cambios en δ_2 y determinar la eficacia relativa de las soluciones aproximadas en comparación con la solución exacta proporcionada por SVD. Este diseño experimental busca comprender el rendimiento del gradiente descendente aplicado a problemas de cuadrados mínimos, considerando tanto la solución estándar como la regularizada mediante L2, y evaluando la robustez frente a variaciones en el parámetro de regularización.

3.2 Optimización de Portfolio de inversión por optimización cuadrática convexa con restricciones

El experimento comenzó con la carga de datos desde el archivo que tenía la información sobre los retornos históricos de 100 empresas en 2 años/24 meses. Se aseguró la numericidad de los datos, para poder trabajar con estos como valores de punto flotante (del tipo `float`). Para evitar sesgos causados por valores de diferentes magnitudes, estos fueron escalados en un rango del 0 al 1.

A continuación, con estos datos "más lindos" sobre los retornos históricos se procedió a calcular la matriz de covarianza con la función `cov` para poder tener información sobre las relaciones entre diferentes activos y también se calculó el vector de retornos promedio con la función `mean` para tener una idea sobre los rendimientos esperados. Ambas funciones utilizadas fueron sacadas de la librería Numpy.

Para poder realizar la optimización de manera correcta se definió la función objetivo que busca minimizar el riesgo y las restricciones a las cuales se tiene que amoldar. Luego se realizó la optimización con el método de SLSQP para encontrar el mix óptimo de los activos que permite la minimización del riesgo, que al mismo tiempo cumple las restricciones. Durante este proceso se mantuvo un registro de los datos en cada iteración para poder luego hacer una comparación de convergencia.

Finalmente se realizó la optimización para diferentes casos con diferentes valores de $r_{\min}$ y se comparó en un histograma los diferentes riesgos asociados a cada mix óptimo de activos. Además, el código devuelve el mix óptimo en consola.

4 Análisis de los resultados

4.1 Experimento 1: Cuadrados mínimos mediante descenso por gradiente

En este experimento, se aplicó el algoritmo de descenso por gradiente para resolver sistemas lineales sobredeterminados. Se exploró la regularización L2 y se compararon los resultados con la solución exacta obtenida mediante la Descomposición de Valores Singulares (SVD).

Análisis de Convergencia

La figura 1 muestra el comportamiento del error a lo largo de las iteraciones utilizando el descenso por gradiente con y sin la regularización L2. Se observa una tendencia descendente en el error a medida que aumentan las iteraciones, indicando la convergencia del algoritmo.

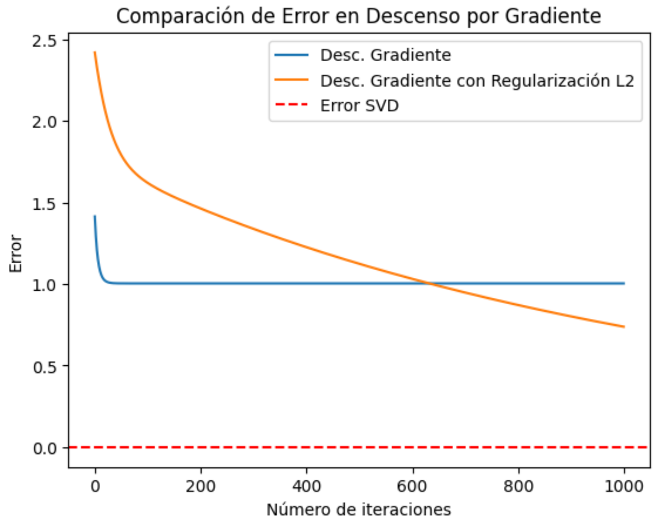

Figure 1: Error con Gradiente Descendente

La figura 2 presenta la variación del error para distintos valores de (δ^2). Se aprecia que a medida que se reduce el valor de δ^2, el error tiende a disminuir en las primeras iteraciones. Sin embargo, para valores muy pequeños de δ^2, se observa un estancamiento o incluso un aumento del error, sugiriendo una posible influencia negativa de una regularización excesivamente baja en la convergencia del algoritmo.

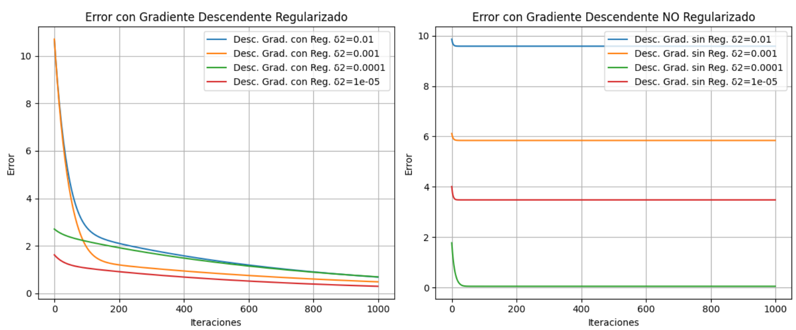

Figure 2: Variación del Error con Diferentes Valores de δ^2

Selección del Parámetro de Tamaño de Paso (s)

En el contexto del algoritmo de Gradiente Descendente, la elección del tamaño del paso (s) es crucial para la convergencia y la eficiencia del método. Idealmente, queremos seleccionar un valor de s que permita que el algoritmo converja rápidamente hacia el mínimo de la función de costo. La elección común es utilizar $s = \frac{1}{\lambda_{\max}}$, donde $\lambda_{\max}$ es el autovalor máximo de la matriz Hessiana de la función de costo. Esta elección tiene en cuenta la curvatura de la función en el punto actual y ajusta el tamaño del paso en consecuencia.

Cuando se utiliza esta fórmula, el algoritmo tiende a converger de manera más eficiente. Sin embargo, la selección de s puede depender del problema específico y, en la práctica, puede requerir ajustes y experimentación para encontrar el valor óptimo. El valor de s se elige para equilibrar la rapidez de convergencia y la estabilidad del algoritmo, y la elección de $s = \frac{1}{\lambda_{\max}}$ es una opción común y efectiva.

Comparación con Solución Exacta mediante SVD

Se compararon las soluciones obtenidas mediante el descenso por gradiente con regularización L2 con la solución exacta proporcionada por la SVD. A pesar de la aproximación, se observó una tendencia a converger hacia valores cercanos a la solución exacta, especialmente con valores adecuados de regularización. Se puede visualizar la diferencia entre las soluciones obtenidas mediante SVD y las soluciones por gradiente descendente (con y sin regularización) se utilizando un gráfico de barras para comparar las normas de esas diferencias en la figura 3. Se puede observar que el al implementar la regularizacion utilizando $\delta^2 = 0.01$ este método realiza una mejor aproximación.

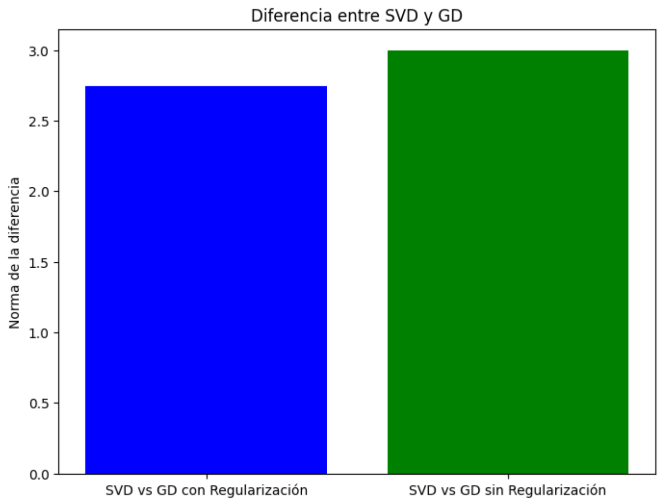

Figure 3: Comapración con solución exacta de SVD

Los experimentos demostraron la capacidad del descenso por gradiente con y sin regularización L2 para resolver sistemas lineales sobredeterminados. La elección adecuada de parámetros, como el tamaño de paso y el valor de regularización, resultó crucial para la convergencia y la precisión de las soluciones aproximadas.

4.2 Experimento 2: Optimización de Portfolio de inversión por optimización cuádratica convexa con restricciones

Se realizó la optimización SLSQP para 9 valores distindos de $r_{\min}$. Estos fueron los siguientes:

$$r_{\min} = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\} \tag{1}$$

Posteriorimente se realizó el histograma en el cual se comparaban los riesgos asociados a cada mix óptimo de los activos. Este fue el resultado de ese ploteo:

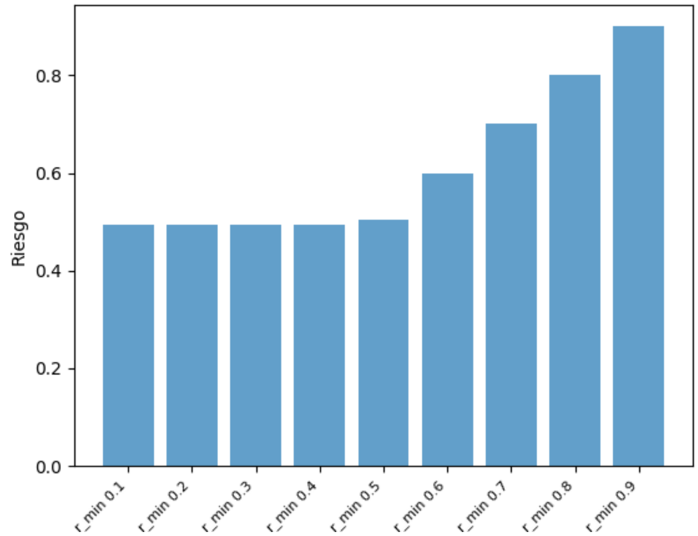

Figure 4: Riesgo en comparación al crecimiento del valor $r_{\min}$ del rendimiento mínimo esperado

A simple vista fue posible observar como el riesgo se mantuvo igual hasta que el valor de $r_{\min}$ creció hasta 0.5, allí el riesgo comenzó a subir linealmente. La restricción $r_{\min}$ del rendimiento mínimo, cuando posee un valor más bajo es más fácil encontrar una combinación de activos que cumpla con esta restricción, ya que hay una variedad de asignaciones que pueden proporcionar el rendimiento mínimo deseado sin incurrir en un riesgo significativo. Sin embargo, a medida que incrementa el valor de $r_{\min}$, la búsqueda de soluciones se vuelve más desafiante; y en consecuencia aumentando el riesgo.
Entonces si resumimos lo que fue analizado lleagmos a que, al imponer un rendimiento mínimo más alto como restricción, el algoritmo de optimización puede verse limitado en la elección de activos y ponderaciones, lo que conduce a soluciones que tienen un mayor riesgo. Y que el riesgo comenzó a subir a partir de que el valor de $r_{\min}$ estuvo entre 0.4 y 0.5, ya que con valores menores se mantuvo casi igual.

5 Concusiones

Tras analizar el rendimiento del algoritmo de Gradiente Descendente en la resolución de sistemas lineales sobredeterminados y explorar la optimización de portafolios de inversión bajo restricciones, se destaca la eficacia de la regularización L2 para mejorar la aproximación de soluciones en ambos contextos. En el primer experimento, la aplicación de la regularización L2 en el Gradiente Descendente condujo a una convergencia más eficiente y precisa, aunque se evidenció la sensibilidad al parámetro de regularización. Asimismo, se observó que en el segundo experimento, que tras analizar el crecimiento del riesgo luego de aumentar el valor de la restricción $r_{\min}$ se pudo llegar a la conclusión de que hay una relación de causalidad entre el aumento de $r_{\min}$ y el riesgo; mostrando que sería óptimo realizar el análisis de inversión con restricciones del rendimiento mínimo esperado más bajas para así minimizar el riesgo.

Los resultados respaldan la utilidad del Gradiente Descendente con regularización y destacan la importancia de las restricciones en la optimización de portafolios para mitigar el riesgo en la toma de decisiones financieras

6 Bibliografía

- Nocedal, J., & Wright, S. J. (Eds.). (1999). *Numerical optimization.* New York, NY: Springer New York.
- Strang, G. (2019). *Linear algebra and learning from data (Vol. 4).* Cambridge: Wellesley-Cambridge Press.
- Burden, R. L., Faires, J. D., & Burden, A. M. (2015). Numerical analysis. Cengage learning.
- Trefethen, L. N., & Bau, D. (2022). Numerical linear algebra (Vol. 181). Siam.