

Práctica 2: Optimización con restricciones (Implementación del ADALINA LASSO)

Víctor Tirado

victortiradoarregui@uma.es

Algoritmos de Búsqueda y Optimización Computacional. Universidad de Málaga.

En esta práctica se aborda la resolución de un problema de regresión lineal con regularización 1 (LASSO), el cual tiene como objetivo evitar el sobreajuste y realizar selección de características automáticamente. Para ello, se utiliza el método de Gradiente Projectado como técnica de optimización.

1 Método del Gradiente Projectado con Regularización ℓ_1

El método del Gradiente Projectado consiste en iterar en la dirección del gradiente, proyectando en cada paso la solución sobre un conjunto factible que representa la restricción de la norma ℓ_1 .

La función objetivo es:

$$\min_{w \in R^K} f(w) = \frac{1}{2} \|Xw - y\|^2 + \lambda \|w\|_1$$

Donde:

- $X \in R^{N \times K}$ es la matriz de características.
- $y \in R^N$ es el vector de valores objetivos.
- $w \in R^K$ es el vector de pesos.
- λ es el parámetro de regularización.

Se utiliza **backtracking** para determinar el tamaño de paso α_k , comenzando desde $\alpha_{init} = 1$ y multiplicando iterativamente por un factor $\beta < 1$ (por ejemplo, $\beta = 0.5$) hasta satisfacer una condición de descenso suficiente. El algoritmo se detiene tras un máximo de 100 iteraciones o cuando la mejora en el error es menor que 10^{-3} .

2 Resultados con $\lambda = 100$

- **Número de iteraciones hasta la convergencia:** 14.
- **Coefficientes aprendidos:** [-0. -0. 0.06042465 0.44253515 0. -0.1467933 0.13784139 -0. -0. 0.00981565 -0.27648231 -0.08349461 0. -0.53595042].

2.1 Comparación de valores reales y predichos para las 5 primeras viviendas

Vivienda	Valor real	Valor predicho
1	24	26.19
2	21.60	26.18
3	34.70	31.09
4	33.40	28.72
5	36.20	27.34

3 Análisis del efecto de λ

Se han probado distintos valores de λ : 10, 50, 100 y 200. A continuación se analiza su efecto sobre la selección de características y el ajuste del modelo:

- $\lambda = 10$: Mínima penalización, la mayoría de los coeficientes no son cero. Mayor riesgo de sobreajuste.
- $\lambda = 50$: Aparecen ceros en los coeficientes. Mejora la generalización.
- $\lambda = 100$: Modelo más simple, muchos coeficientes se anulan. Buen equilibrio entre ajuste y simplicidad.
- $\lambda = 200$: Fuerte regularización, pocos coeficientes distintos de cero. Posible subajuste.

Gráfica del número de características seleccionadas ($w_i \neq 0$) frente a λ .

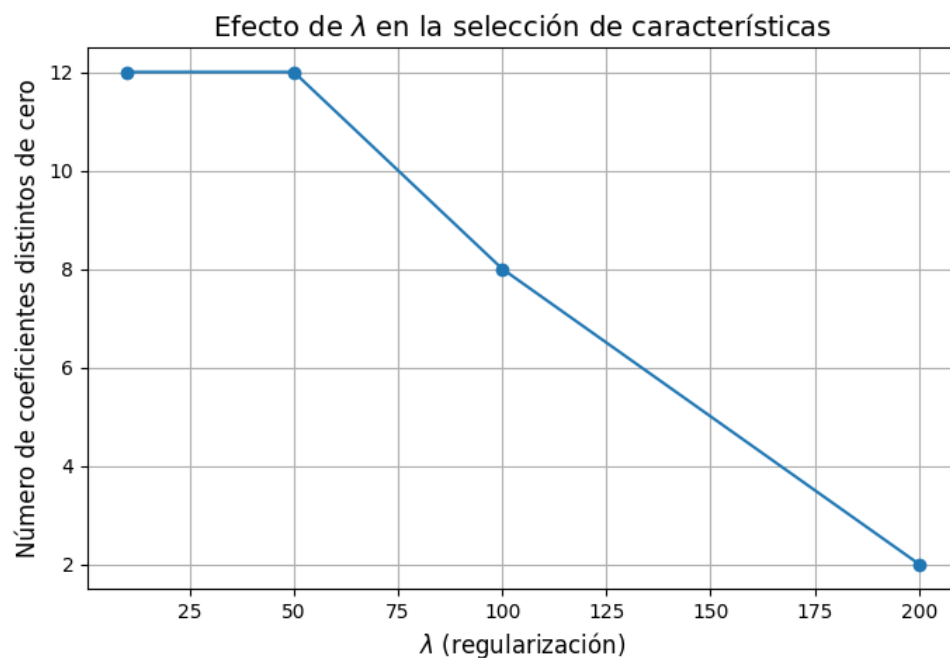


Fig. 1. Relación entre el valor de λ y el número de características seleccionadas ($w_i \neq 0$).

4 Conclusiones

- El método de Gradiente Proyectado permite resolver eficazmente problemas de optimización con restricciones como el LASSO.
- La regularización ℓ_1 introduce sparsity en el modelo, eliminando características irrelevantes.
- El parámetro λ debe ser cuidadosamente seleccionado para equilibrar ajuste y complejidad.