

# Optimización de Hiperparámetros en Métodos de Regresión Penalizada para Alta Dimensionalidad

Edison Antony Sayritupa Coaricona

## Introducción

La regresión penalizada se ha convertido en una técnica fundamental en el análisis de datos cuando el número de predictores es mayor que el número de observaciones. Métodos como **Lasso**, **Ridge** y **Elastic Net** son particularmente efectivos para manejar alta dimensionalidad y prevenir el sobreajuste en estos contextos.

La introducción de términos de penalización en la función de costo permite controlar la complejidad del modelo y mejorar su capacidad de generalización. En particular:

- La penalización  $\ell_1$  en **Lasso** y la combinación  $\ell_1$ - $\ell_2$  en **Elastic Net** permiten la selección de variables.
- **Ridge** aplica una penalización  $\ell_2$  para reducir la magnitud de los coeficientes sin eliminar predictores.

La capacidad de manejar datos altamente correlacionados y generar soluciones escasas hacen que estos métodos sean opciones atractivas en problemas de regresión con grandes conjuntos de datos.

El principal desafío en la aplicación de estos métodos radica en la correcta optimización de los parámetros de penalización, específicamente el valor de  $\lambda$ , que regula la intensidad de la penalización y tiene un gran impacto en la precisión del modelo.

La selección de  $\lambda$  puede lograrse mediante técnicas de validación cruzada, que proporcionan una estimación robusta del desempeño del modelo en datos no observados. Además, se utilizan enfoques de búsqueda automatizada, como la búsqueda en cuadrícula y la optimización bayesiana, para identificar los valores óptimos de los hiperparámetros.

Estos enfoques no solo mejoran la capacidad predictiva del modelo, sino que también permiten una mejor interpretación de los resultados y mayor estabilidad en presencia de datos ruidosos o correlacionados.

En este trabajo se presenta una revisión de las estrategias más utilizadas y se discute su aplicabilidad en contextos reales. En particular, se analizan los efectos de la optimización de  $\lambda$  sobre el desempeño e interpretabilidad de los modelos de regresión penalizada en conjuntos de datos de alta dimensionalidad, destacando los avances más recientes en técnicas de regularización y selección de variables.