

# Outros modelos de regressão linear

## Regressão Ridge

Ridge regression é uma técnica de regularização utilizada em regressão linear para lidar com problemas de colinearidade. Ele adiciona uma penalidade L2 à função de custo, o que faz com que os coeficientes sejam menores. Isso ajuda a evitar o overfitting, já que modelos com coeficientes muito altos tendem a se ajustar a ruídos nos dados de treinamento, em vez de generalizar bem para dados desconhecidos. Em resumo, o Ridge Regression é uma técnica que ajuda a manter os coeficientes do modelo sob controle, melhorando a capacidade de generalização do modelo.

A fórmula para calcular o erro quadrático da Ridge Regression é dada por:

$$SSE = \sum_{i=1}^n (reg_i - y_i)^2 + \lambda \sum_{i=1}^n (m_i)^2, \text{ onde } reg = m_1x_1 + m_2x_2 + \dots + m_nx_n + b$$

O hiperparâmetro lambda é usado para controlar a magnitude da penalidade. Quanto maior o lambda, maior a regularização e menores os coeficientes. Um lambda muito grande pode levar a underfitting, enquanto um lambda muito pequeno não terá efeito suficiente na regularização.

## Regressão Lasso

A regressão Lasso é uma técnica de regularização de regressão linear que adiciona uma penalidade L1 ao erro de treinamento. Isso significa que ao invés de apenas minimizar o erro quadrático entre as previsões e os valores alvo, a regressão Lasso também minimiza a soma dos valores absolutos dos coeficientes (chamados de pesos) dos modelos. Isso tende a tornar alguns dos coeficientes zero, o que se traduz em seleção automática de recursos. Isso é útil quando há muitos recursos e alguns deles são irrelevantes ou redundantes.

A fórmula para calcular o erro quadrático da Lasso Regression é dada por:

$$SSE = \sum_{i=1}^n (reg_i - y_i)^2 + \lambda \sum_{i=1}^n |m_i|, \text{ onde } reg = m_1x_1 + m_2x_2 + \dots + m_nx_n + b$$

Lambda é o parâmetro de regularização, que controla o grau de penalização dos coeficientes. Quanto maior o valor de  $\lambda$ , maior será a penalização e maior será a probabilidade de alguns dos coeficientes se tornarem zero.

## Regressão Elastic Net

A regressão Elastic Net é uma técnica de regularização que combina as características da regressão Ridge e Lasso. Ela adiciona uma penalidade L1 e L2 ao erro de treinamento, o

que significa que ela minimiza tanto a soma dos quadrados dos coeficientes quanto a soma dos valores absolutos dos coeficientes. Isso tende a tornar alguns dos coeficientes zero (como na Lasso) e distribuir os valores dos coeficientes de forma mais homogênea entre os recursos relevantes e irrelevantes (como na Ridge).

A fórmula para calcular o erro quadrático da Lasso Regression é dada por:

$$SSE = \sum_{i=1}^n (reg_i - y_i)^2 + (\lambda \times c) \times \sum_{i=1}^n |m_i| + \lambda \times (1 - c) \times \sum_{i=1}^n (m_i)^2, \text{ onde}$$

$$reg = m_1x_1 + m_2x_2 + \dots + m_nx_n + b$$

Lambda continua sendo o parâmetro de regularização, que controla o grau de penalização dos coeficientes. Quanto maior o valor de  $\lambda$ , maior será a penalização.

A constante  $c$  é o parâmetro que controla a proporção entre a penalidade L1 e L2, sendo ela  $0 \leq c \leq 1$ . Quanto maior o valor de  $c$ , maior será a contribuição da penalidade L1 e menor será a contribuição da penalidade L2.