

# Linear Regression with Ridge Regularization

## Problem Statement

Consider a dataset where we aim to predict an outcome based on a set of input features. However, we want to avoid overfitting by controlling the complexity of the model. The objective is to formulate an optimization problem that balances the trade-off between fitting the training data well and preventing overfitting caused by excessive model complexity.

## Mathematical Formulation

Let  $\mathbf{x}$  denote the input vector and  $\mathbf{y}$  represent the associated target value corresponding to the input vector  $\mathbf{x}$  in the training dataset. Assuming a linear relationship, the predicted value  $\hat{\mathbf{y}}$  for input vector  $\mathbf{x}$  is given by  $\hat{\mathbf{y}} = \mathbf{w}^T \mathbf{x} + \mathbf{b}$ , where  $\mathbf{w}$  is the coefficient vector and  $\mathbf{b}$  is the bias vector.

The error between the predicted value  $\hat{\mathbf{y}}$  and the actual target value  $\mathbf{y}$  for input vector  $\mathbf{x}$  can be quantified using the squared difference:

$$\text{error} = \|\mathbf{w}^T \mathbf{x} + \mathbf{b} - \mathbf{y}\|^2$$

The goal is to minimize this error to achieve the best fit for the training data. However, overfitting can occur when the model becomes too complex, leading to poor generalization to unseen data. To address overfitting, Ridge regression is employed, which adds a regularization term to the objective function. The regularized objective function becomes:

$$\text{minimize } \|\mathbf{w}^T \mathbf{x} + \mathbf{b} - \mathbf{y}\|^2 + \lambda \|\mathbf{w}\|^2$$

where  $\lambda$  is the regularization parameter that controls the strength of the penalty. The regularization term  $\|\mathbf{w}\|^2$  penalizes large coefficients in the coefficient vector  $\mathbf{w}$ , thereby encouraging simpler models that generalize better to unseen data.

## Summary

Ridge regularization is a powerful technique in linear regression for preventing overfitting by penalizing large coefficients. By incorporating a regularization

term into the objective function, Ridge regression promotes models that strike a balance between fitting the training data well and controlling model complexity. This results in more robust models that generalize effectively to unseen data.