L1 Regularization

Shah Md. Arshad Rahman Ziban

1   Introduction

Regularization is a technique used in machine learning to prevent overfitting by adding a penalty term to the loss function. This ensures that the model generalizes well to new, unseen data. One common type of regularization is L1 regularization, also known as Lasso regularization.

2   L1 Regularization (Lasso)

L1 regularization adds a penalty based on the absolute values of the model's weights. This has the effect of making some of the weights exactly zero, meaning some features are completely removed. This helps in feature selection by identifying and retaining only the most important features.

- Effect: It makes some of the weights exactly zero, effectively removing some features.

- Use Case: Use L1 regularization when you want to perform feature selection and remove unnecessary features.

Example

Imagine you're choosing ingredients for a recipe. L1 regularization removes the unnecessary ingredients, leaving only the most important ones.

3   Lasso Regression Steps

To predict a value using Lasso regression, follow these steps:

1. Generate the dataset.

2. Fit the Lasso regression model to the dataset.

3. Use the model to predict the value for the given input.

Step 1: Generate the dataset

We need to generate random values for X and y. Since the exact values are not provided, we will generate them using Python's numpy library.

```python
import numpy as np

# Generate random values for X and y
X = np.random.randint(1, 100, size2=0)
y = np.random.randint(1, 100, size=20)
```

Step 2: Fit the Lasso regression model

We will use the Lasso class from the sklearn.linear $_m$ odelmoduletofitthemodel.

```python
from sklearn.linear_model import Lasso

# Reshape X to a 2D array
X = X.reshape(-1, 1)

# Initialize the Lasso regression model
lasso = Lasso(alpha=1.0)

# Fit the model to the data
lasso.fit(X, y)
```

Why Reshape?

When you generate X using np.random.randint(1, 100, size=20), it creates a 1D array of 20 random integers. However, scikit-learn's Lasso class expects X to be a 2D array. Here's why:

- Input Format Requirement: Scikit-learn's fit method expects X to be a 2D array where each row is a sample and each column is a feature.

- Consistency: Reshaping ensures that the data is in the correct format for the algorithm to process it correctly.

Step 3: Predict the value for the given input

We need to predict the value for the input 103.

```python
# Predict the value for X = 103
prediction = lasso.predict([[103]])

# Print the prediction
print("The predicted value for 103 is : ", prediction[0])
```

4 Lasso Regression Overview

Lasso Regression is a type oflinear regression that includes an L1 regularization term to penalize large coefficients. This helps in reducing overfitting and performing feature selection. Essentially, Lasso regression is a specialized form of linear regression that incorporates regularization to improve model performance and interpretability.

5 Conclusion

The predicted value for 103 using Lasso regression is $\boxed{50}$.

Note: The actual predicted value may vary each time you run the code because the dataset is randomly generated.