

Ridge Regression vs. Lasso Regression

1. Importing necessary libraries

First, we need to import the libraries required for our analysis. We will use `numpy` for numerical operations, `pandas` for data manipulation, `matplotlib` for visualisation, and `sklearn` for loading the dataset and performing regression.

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.datasets import fetch_california_housing
from sklearn.linear_model import Ridge, Lasso
from sklearn.model_selection import train_test_split
```

2. Loading the dataset

We load the California housing prices dataset from `sklearn`. For the sake of simplicity, we will focus on only one feature (`MedInc`, i.e., median income) along with the target variable (housing price) for our regression analysis. We also split the dataset into training and testing sets for evaluating our models.

```
california_housing = fetch_california_housing()
X = california_housing.data
y = california_housing.target

X = X[:, 0].reshape(-1,1)    # reshaping to a 2D array

X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

3. Implementing ridge regression

Now, we fit ridge regression models to our training data with different values for the tuning parameter (λ), and use these to make predictions on our testing set.

```
ridge_models = {}
lambdas = [0.1, 1.0, 10.0]

for l in lambdas:
    ridge_model = Ridge(alpha=l)
    ridge_model.fit(X_train, y_train)
    ridge_models[l] = ridge_model.predict(X_test)
```

4. Implementing lasso regression

Next, we fit lasso regression models to our training data with the same lambda values as ridge regression, and use these to make predictions on our testing set.

```
lasso_models = {}
lambdas = [0.1, 1.0, 10.0]

for l in lambdas:
    lasso_model = Lasso(alpha=l)
    lasso_model.fit(X_train, y_train)
    lasso_models[l] = lasso_model.predict(X_test)
```

5. Visualising the results

Finally, we visualise the results by plotting the test data, as well as the regression lines for both ridge and lasso regression. This will help us compare the two regression methods.

```
plt.figure(figsize=(12,6))

plt.scatter(X_test, y_test, color='black', label='Data')

for l, y_pred_ridge in ridge_models.items():
    plt.plot(X_test, y_pred_ridge, linewidth=2, label=f'Ridge regression ( $\lambda={l}$ )')

for l, y_pred_lasso in lasso_models.items():
    plt.plot(X_test, y_pred_lasso, linestyle='--', linewidth=2, label=f'Lasso regression ( $\lambda={l}$ )')

plt.xlabel('MedInc')
plt.ylabel('House Price')
```

```
plt.title('Comparison of Ridge and Lasso Regression')
plt.xlabel('Median Income (MedInc)')
plt.ylabel('Housing Price')
plt.legend()
plt.show()
```

