Project 1 on Machine Learning

Maziar Kosarifar

October 2019

Abstract

In this project we aim to evaluate and study different features, and properties of three different regression methods including Ordinary Least Squares(OLS), Ridge regression and Lasso regression. And the application of resampling techniques. And we compare their result for Franke function and then introduce real terrain data, as inputs.

Introduction

Regression methods are widely used to study the relation between two variables, and estimate if and how they are related, and understand whether there is a close numerical correlation between variables. Ordinary Least Squares, Ridge regression and Lasso regression are the one implemented in this project.

In this paper I will briefly explain each one of them. Then I'll introduce the use of statistical methods to evaluate each of these regression techniques.

In the last part we will introduce real world data of geographical terrains as the input, to evaluate the differences between these methods.

Ordinary Least Squares

Ordinary Least Squares are one of the methods used in approximation theory, which seeks to find an approximate curve that would minimize the error norm, when given a set of real values such as f_0, f_1, ..., f_n at real data points X_0, X_1, ..., X_n which we call nodes.

So we want to find a polynomial P(X):

$$P(x, y) = \Sigma_{i=0}^{n} \beta_i x^i y^{n-i}$$

to approximate the value of F(x, y).

This problem can be turned into:

$$y = X\beta + \epsilon \tag{1}$$

This can be turned into a problem of minimizing the following value of S

$$S = ||X\beta - y||_2^2 \tag{2}$$

Which is a set of normal equations, and the minimum value will be found when the gradient of S is zero, so we have:

$$X^T X\beta = X^T y$$

Which can be solved with:

$$\beta = (X^T X)^{-1} X^T y \tag{3}$$

```
def OLS(X, expected_value):
    beta = np.linalg.inv(X.T.dot(X)).dot(X.T).dot(expected_value)
    return beta
```

Ridge regression

OLS treats all variables as equals, and considers that the result equally depends on all of them. Ridge regression on the other hand introduces λ to add some bias to the estimation.

$$\beta^{Ridge} = (X^T X + \lambda I)^{-1} X^T y \tag{4}$$

```python
def OLS(X, expected_value):
    beta = np.linalg.inv(X.T.dot(X)).dot(X.T).dot(expected_value)
    return beta
```