# Types of Regression

SOHAM PATEL (46)

My blog link : [click here](https://rovae.in/digging-ridge-and-lasso-regression/)

Regression techniques are mainly used to identify the relationships among different variables. Its main goal is to predict the value of the dependent variable on the basis of given independent variables.

Regression refers to a line or a curve that passes through all the datapoints on the X-Y plot in such a way that the vertical distance between the line and the datapoints is minimum. This distance measures the relationship, know as correlation. Thus, a ‘best-fit’ model is the one that has a strong relationship.

Most common Regression algorithms used in Machine Learning are **Linear and Logistic Regression.** But, in this discussion we are going to cover **Ridge and LASSO Regression**, which are later version of Simple Linear Regression.

## Ridge Regression

The first question that we should think is “ Why do we need Ridge Regression when we already have Linear Regression? “. Sometimes the Linear Regression model becomes so complex due to high collinearity among the independent variable that the model start overfitting. Hence to save the model from being overfitted, a penalty parameter is added to the regression calculations that yield a reduction in standard errors through ridge regression.

The method of adding a new parameter(lets say α) to the cost function inorder to avoid overfitting, is termed as **L2 regularization**. Here, we penalize the θ parameter by adjusting the value of α in cost function .The cost function for ridge regression can be represented as :

**J( θ ) = Min(||Y – X( θ )||^2 + α||θ||^2)**

where Y = Output feature

X = Input feature

θ = Weights

α = L2 Reguralization parameter

By changing the values of α, we are controlling the penalty term. Higher the values of α, bigger is the penalty and therefore the magnitude of θ is shrinked. Hence, our objective of preventing overfitting is accomplished as the collinearity is reduced along with θ.

## LASSO Regression

LASSO (Least Absolute Shrinkage Selector Operator) is the advance version of Ridge Regression. One additional task that the LASSO Regression performs is feature selection.

Wherever there is strong collinearity amony any two of the feature, LASSO Regression directly drop any one feature by making its coeffcient close to 0, which is not done by Ridge Regression. This method is called **L1 regularization**.

The cost function for LASSO regression can be represented as :

**J( θ ) = Min(||Y – X( θ )||^2 + α||θ||)**

where Y = Output feature

X = Input feature

θ = Weights

α = L1 Reguralization parameter

The only difference between Ridge and LASSO cost function is that instead of taking the square of the coefficients, magnitudes are taken into account for LASSO. This type of regularization (L1) can lead to zero coefficients i.e. some of the features are completely neglected for the evaluation of output. So Lasso regression not only helps in reducing over-fitting but it can help us in feature selection.

## Elastic Net Regression

Elastic net is basically a combination of both Ridge and Lasso Regression. So it uses both L1 and L2 penality term, therefore its equation look like as follows:

**J( θ ) = Min(||Y – X( θ )||^2 + α1||θ|| + α2|θ||^2)**

where Y = Output feature

X = Input feature

θ = Weights

α1 and α2 = L1 and L2 Reguralization parameter

**Final conclusion :**

I can conclude that whenever one is unaware of the wieghtage of features in the dataset, he/she can use Riddge regression where the weightage of the features would be automatically adjusted. But in the case where one doesnt know, if a feature is to be rejected or used, in that case LASSO regression would be usefull to completely remove any feature unlike Riddge. While Elastic Net is perfect to use which handles everything internally but it is bit slower than the other two.