**Ridge and Lasso Regression in Python**

## **1. Brief Overview**

Ridge and Lasso [regression](https://courses.analyticsvidhya.com/courses/introduction-to-data-science-2?utm_source=blog&utm_medium=RideandLassoRegressionarticle)are powerful techniques generally used for creating parsimonious models in presence of a ‘large’ number of features. Here ‘large’ can typically mean either of two things:

1. Large enough to enhance the tendency of a model to overfit (as low as 10 variables might cause overfitting)
2. Large enough to cause computational challenges. With modern systems, this situation might arise in case of millions or billions of features

Though Ridge and Lasso might appear to work towards a common goal, the inherent properties and practical use cases differ substantially. If you’ve heard of them before, you must know that they work by penalizing the magnitude of coefficients of features along with minimizing the error between predicted and actual observations. These are called ‘regularization’ techniques. The key difference is in how they assign penalty to the coefficients:

1. **Ridge Regression:**
   * Performs L2 regularization, i.e. adds penalty equivalent to **square of the magnitude** of coefficients
   * Minimization objective = LS Obj + α \* (sum of square of coefficients)
2. **Lasso Regression:**
   * Performs L1 regularization, i.e. adds penalty equivalent to **absolute value of the magnitude** of coefficients
   * Minimization objective = LS Obj + α \* (sum of absolute value of coefficients)

Note that here ‘LS Obj’ refers to ‘least squares objective’, i.e. the linear regression objective without regularization.

If terms like ‘penalty’ and ‘regularization’ seem very unfamiliar to you, don’t worry we’ll talk about these in more detail through the course of this article. Before digging further into how they work, lets try to get some intuition into why penalizing the magnitude of coefficients should work in the first place.

## **2. Why Penalize the Magnitude of Coefficients?**

Lets try to understand the impact of model complexity on the magnitude of coefficients. As an example, I have simulated a **sine curve** (between 60° and 300°) and added some random noise using the following code:

## **Introduction**

When we talk about Regression, we often end up discussing [Linear and Logistic Regression](https://courses.analyticsvidhya.com/courses/introduction-to-data-science-2?utm_source=blog&utm_medium=RideandLassoRegressionarticle). But, that’s not the end. Do you know there are [7 types of](https://www.analyticsvidhya.com/blog/2015/08/comprehensive-guide-regression/?utm_source=blog&utm_medium=RideandLassoRegressionarticle)Regressions?

Linear and logistic regression is just the most loved members from the family of regressions.  Last week, I saw a recorded talk at NYC Data Science Academy from **Owen Zhang**, Chief Product Officer at DataRobot. He said, **‘if you are using regression without regularization, you have to be very special!’**. I hope you get what a person of his stature referred to.

I understood it very well and decided to explore regularization techniques in detail.

In this article, I have explained the complex science behind ‘Ridge Regression‘ and ‘Lasso Regression‘ which are the most fundamental regularization techniques used in [data science](https://courses.analyticsvidhya.com/courses/introduction-to-data-science-2?utm_source=blog&utm_medium=RideandLassoRegressionarticle), sadly still not used by many.

The overall idea of regression remains the same. It’s the way in which the model coefficients are determined which makes all the difference. I strongly encourage you to go through multiple regression before reading this. You can take help from [this article](https://www.analyticsvidhya.com/blog/2015/10/regression-python-beginners/?utm_source=blog&utm_medium=RideandLassoRegressionarticle) or any other preferred material.

## **Table of Contents**

1. [Brief Overview](https://www.analyticsvidhya.com/blog/2016/01/ridge-lasso-regression-python-complete-tutorial/#one)– How is Ridge and Lasso Regression different?
2. [Why Penalize the Magnitude of Coefficients](https://www.analyticsvidhya.com/blog/2016/01/ridge-lasso-regression-python-complete-tutorial/#two)– Why should they work?
3. [Ridge Regression](https://www.analyticsvidhya.com/blog/2016/01/ridge-lasso-regression-python-complete-tutorial/#three)– How ridge works?
4. [Lasso Regression](https://www.analyticsvidhya.com/blog/2016/01/ridge-lasso-regression-python-complete-tutorial/#four)– How lasso works?
5. [Sneak Peak into Mathematics (Optional)](https://www.analyticsvidhya.com/blog/2016/01/ridge-lasso-regression-python-complete-tutorial/#five)– Some underlying mathematical principles.
6. [Conclusion](https://www.analyticsvidhya.com/blog/2016/01/ridge-lasso-regression-python-complete-tutorial/#six)– Comparing Ridge and Lasso Regression