



[Home](#)



Prashant Sharma — Updated On July 17th, 2023

[Beginner](#) [Machine Learning](#) [Regression](#)

🔖 Help me

C |

This article was published as a part of the [Data Science Blogathon](#).

Introduction

Regression problems are prevalent in machine learning, and regression analysis is the most often used technique for solving them. It is based on data modelling and entails determining the best fit line that passes through all data points with the shortest distance possible between the line and each data point. While there are other techniques for regression analysis, linear and logistic regression are the most widely used. Ultimately, the type of regression analysis model we adopt will be determined by the nature of the data.

Let us learn more about regression analysis and the various forms of regression models.

Table of contents

- [Introduction](#)
- [What is a Regression Model/Analysis?](#)
- [What is the purpose of a regression model?](#)
- [Types of Regression Models](#)
- [1. Linear Regression](#)
- [2. Logistic Regression](#)
- [3. Polynomial Regression](#)
- [4. Ridge Regression](#)
- [5. Lasso Regression](#)
- [6. Quantile Regression](#)
- [7. Bayesian Linear Regression](#)
- [8. Principal Components Regression](#)
- [9. Partial Least Squares Regression](#)
- [10. Elastic Net Regression](#)
- [Frequently Asked Questions](#)
- [Summary](#)

a dataset's dependent (goal) and independent variables. It is widely used when the dependent and independent variables are linked in a linear or non-linear fashion, and the target variable has a set of continuous values. Thus, regression analysis approaches help establish causal relationships between variables, modelling time series, and forecasting. Regression analysis, for example, is the best way to examine the relationship between sales and advertising expenditures for a corporation.

What is the purpose of a regression model?

Regression analysis is used for one of two purposes: predicting the value of the dependent variable when information about the independent variables is known or predicting the effect of an independent variable on the dependent variable.

Types of Regression Models

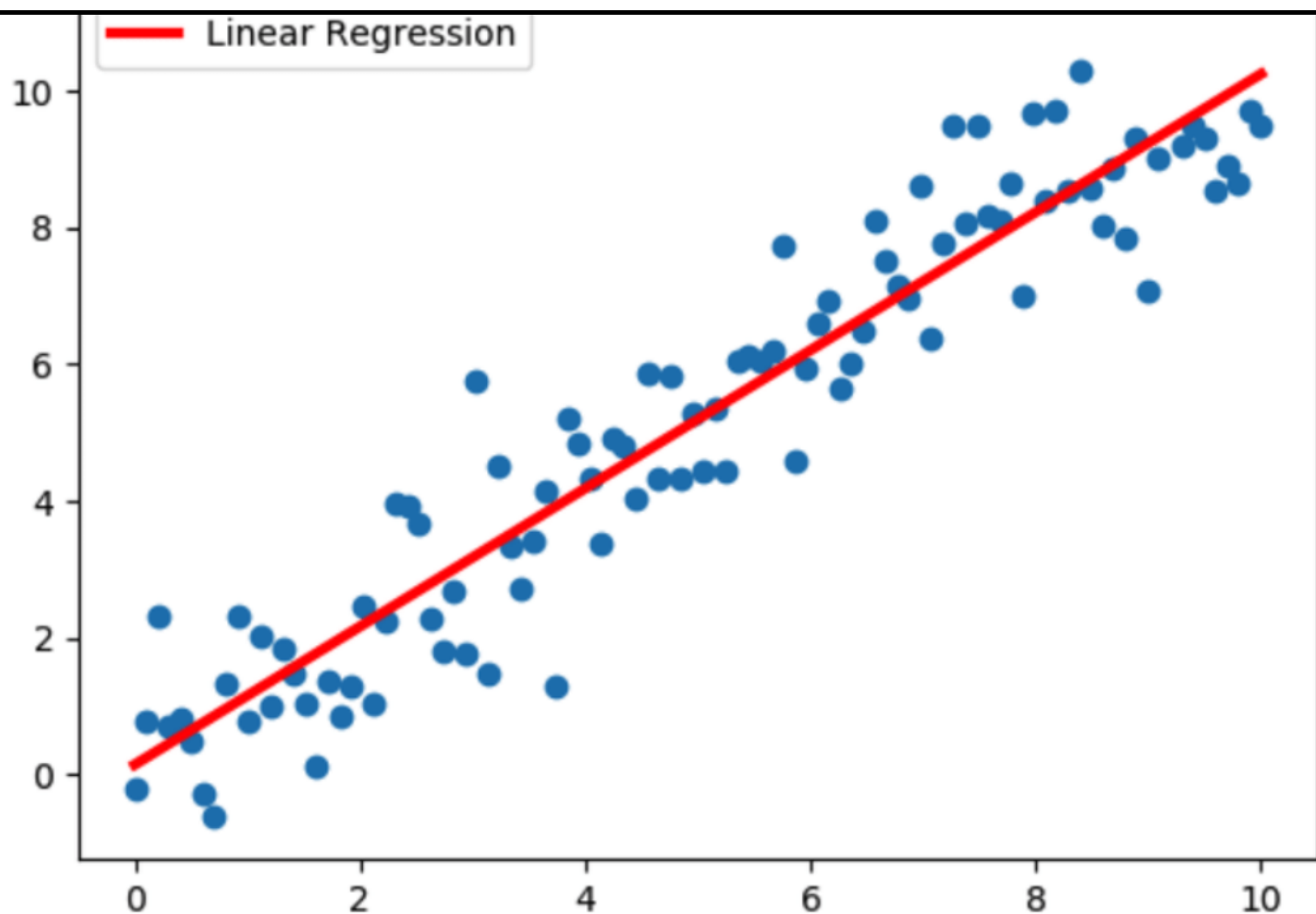
There are numerous regression analysis approaches available for making predictions. Additionally, the choice of technique is determined by various parameters, including the number of independent variables, the form of the regression line, and the type of dependent variable.

Let us examine several of the most often utilized regression analysis techniques:

1. Linear Regression

The most extensively used modelling technique is linear regression, which assumes a linear connection between a dependent variable (Y) and an independent variable (X). It employs a regression line, also known as a best-fit line. The linear connection is defined as $Y = c + m \cdot X + e$, where 'c' denotes the intercept, 'm' denotes the slope of the line, and 'e' is the error term.

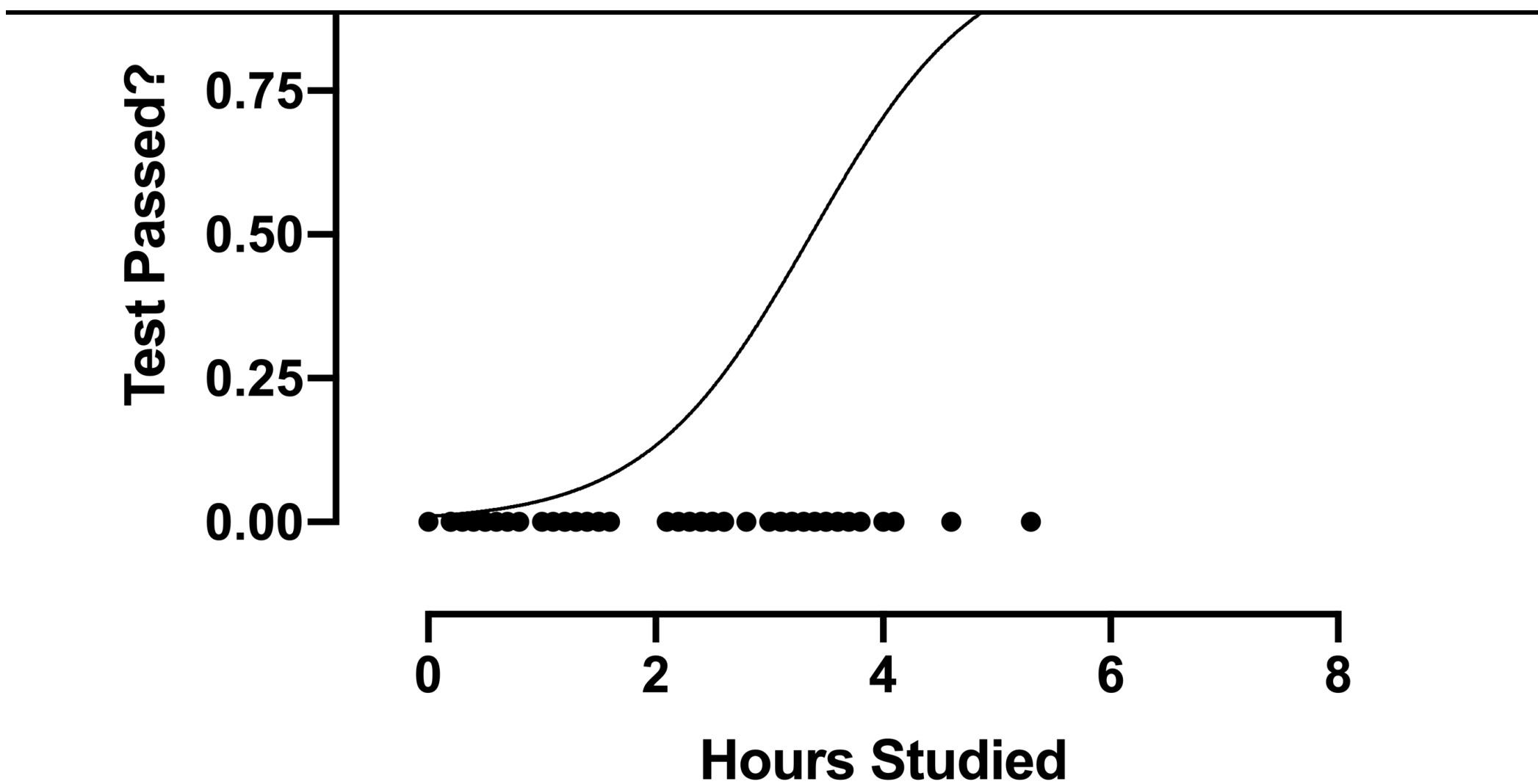
The linear regression model can be simple (with only one dependent and one independent variable) or complex (with numerous dependent and independent variables) (with one dependent variable and more than one independent variable).



[IMAGE](#)

2. Logistic Regression

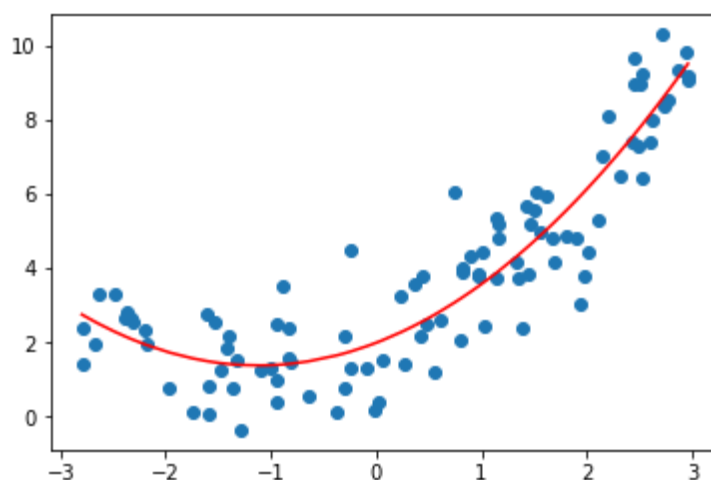
When the dependent variable is discrete, the logistic regression technique is applicable. In other words, this technique is used to compute the probability of mutually exclusive occurrences such as pass/fail, true/false, 0/1, and so forth. Thus, the target variable can take on only one of two values, and a sigmoid curve represents its connection to the independent variable, and probability has a value between 0 and 1.



[IMAGE](#)

3. Polynomial Regression

The technique of polynomial regression analysis is used to represent a non-linear relationship between dependent and independent variables. It is a variant of the multiple linear regression model, except that the best fit line is curved rather than straight.



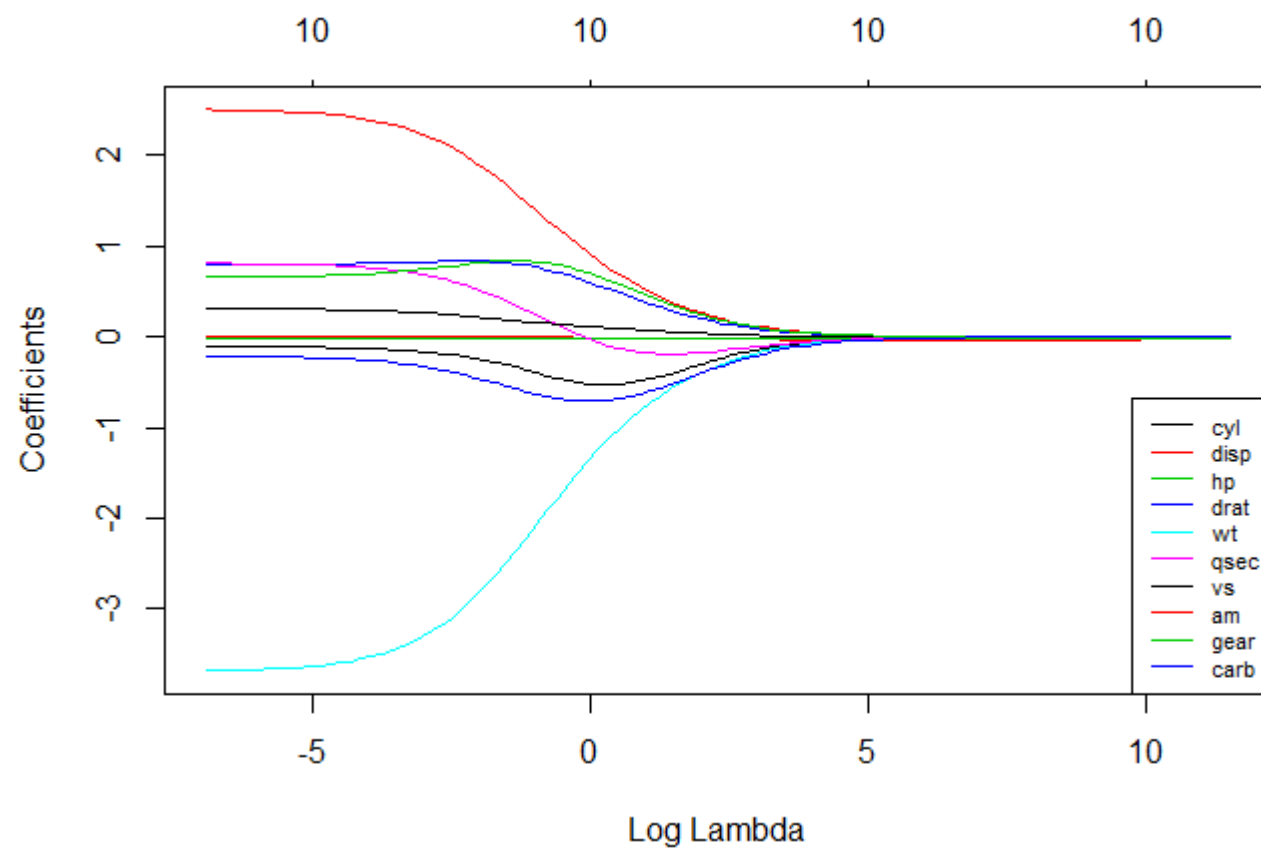
[IMAGE](#)

4. Ridge Regression

When data exhibits multicollinearity, that is, the ridge regression technique is applied when the independent variables are highly correlated. While least squares estimates are unbiased in multicollinearity, their variances are significant enough to cause the observed value to diverge from the actual value. Ridge regression reduces standard errors by biasing the regression estimates.

The lambda (λ) variable in the ridge regression equation resolves the multicollinearity problem.

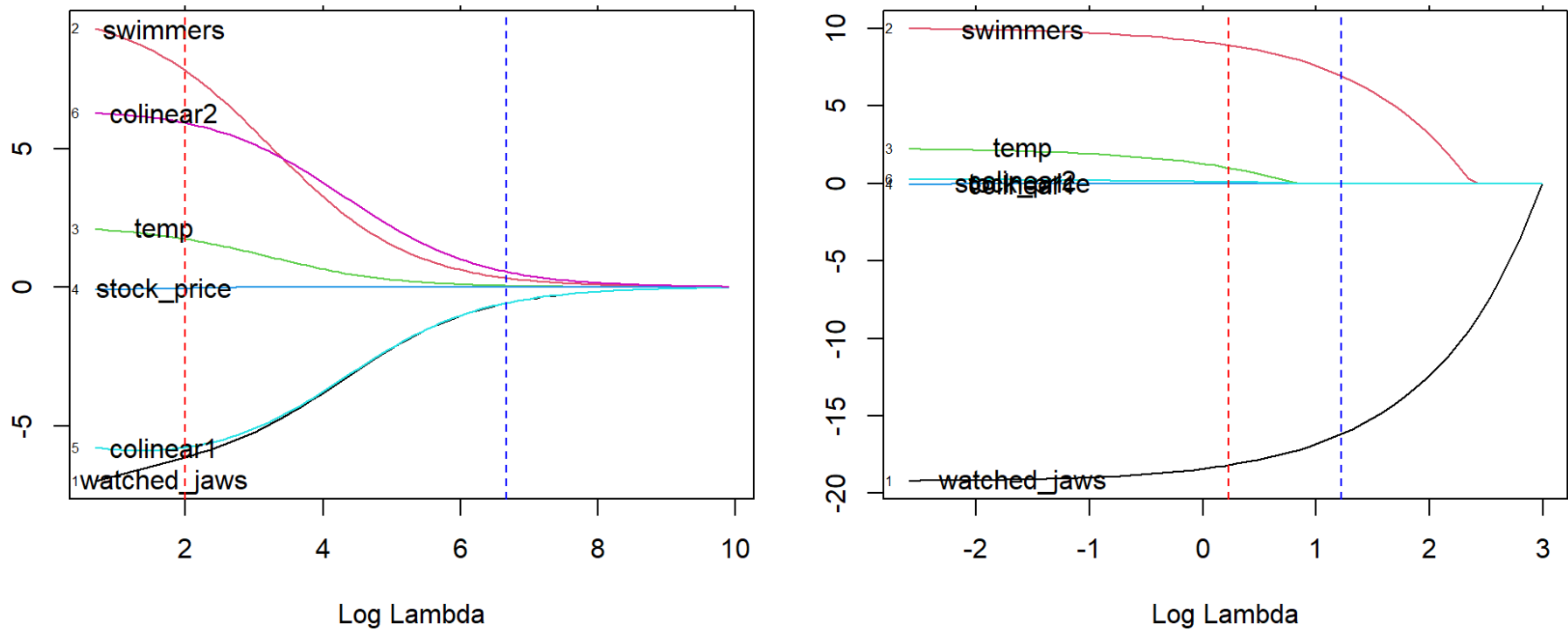
$$= \operatorname{argmin}_{\beta \in \mathbb{R}^p} \underbrace{\|y - X\beta\|_2^2}_{\text{Loss}} + \lambda \underbrace{\|\beta\|_2^2}_{\text{Penalty}}$$



[IMAGE](#)

5. Lasso Regression

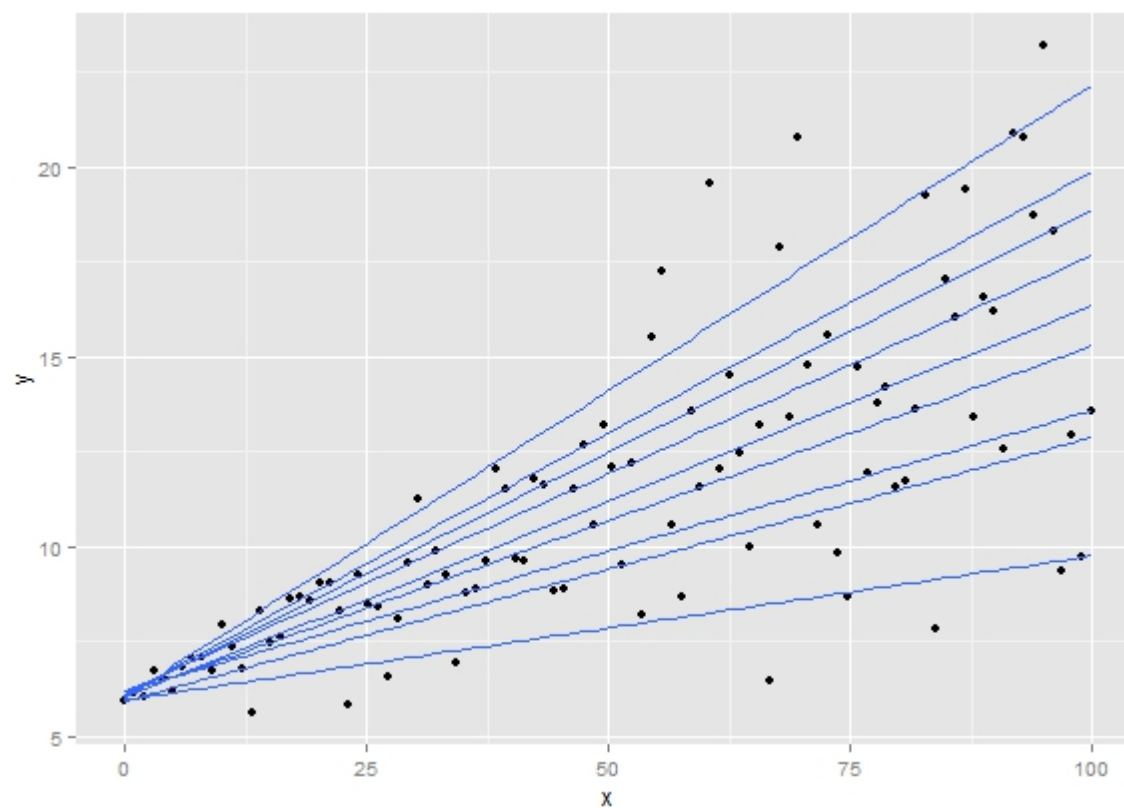
As with ridge regression, the lasso (Least Absolute Shrinkage and Selection Operator) technique penalizes the absolute magnitude of the regression coefficient. Additionally, the lasso regression technique employs variable selection, which leads to the shrinkage of coefficient values to absolute zero.



[IMAGE](#)

6. Quantile Regression

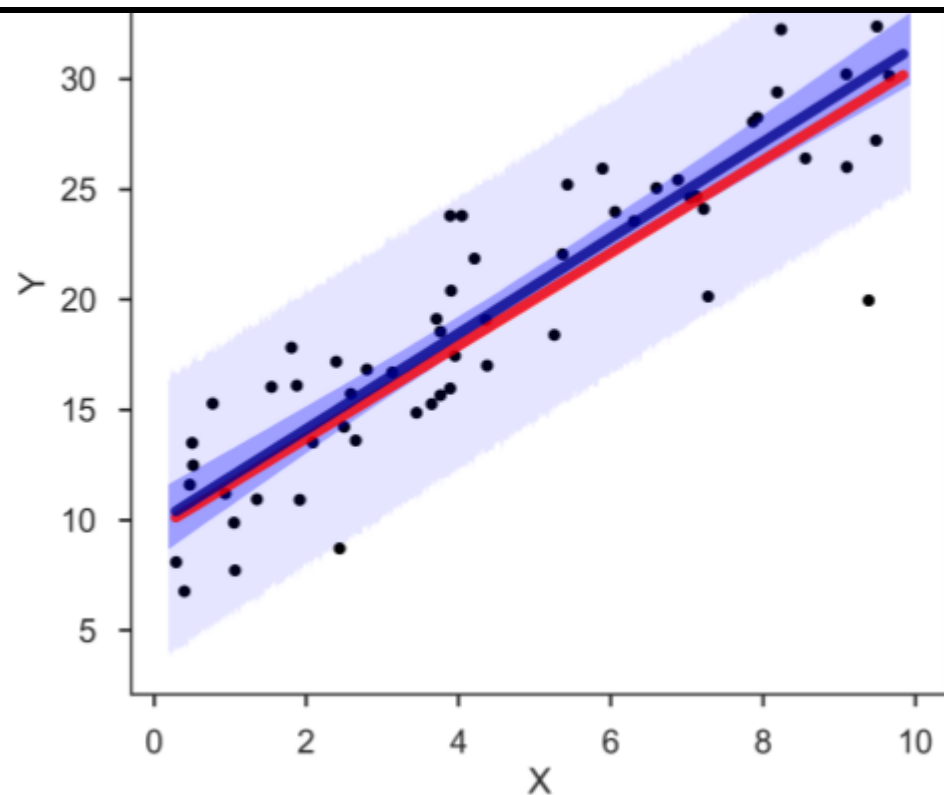
The quantile regression approach is a subset of the linear regression technique. It is employed when the linear regression requirements are not met or when the data contains outliers. In statistics and econometrics, quantile regression is used.



[IMAGE](#)

7. Bayesian Linear Regression

Bayesian linear regression is a form of regression analysis technique used in machine learning that uses Bayes' theorem to calculate the regression coefficients' values. Rather than determining the least-squares, this technique determines the features' posterior distribution. As a result, the approach outperforms ordinary linear regression in terms of stability.



[IMAGE](#)

8. Principal Components Regression

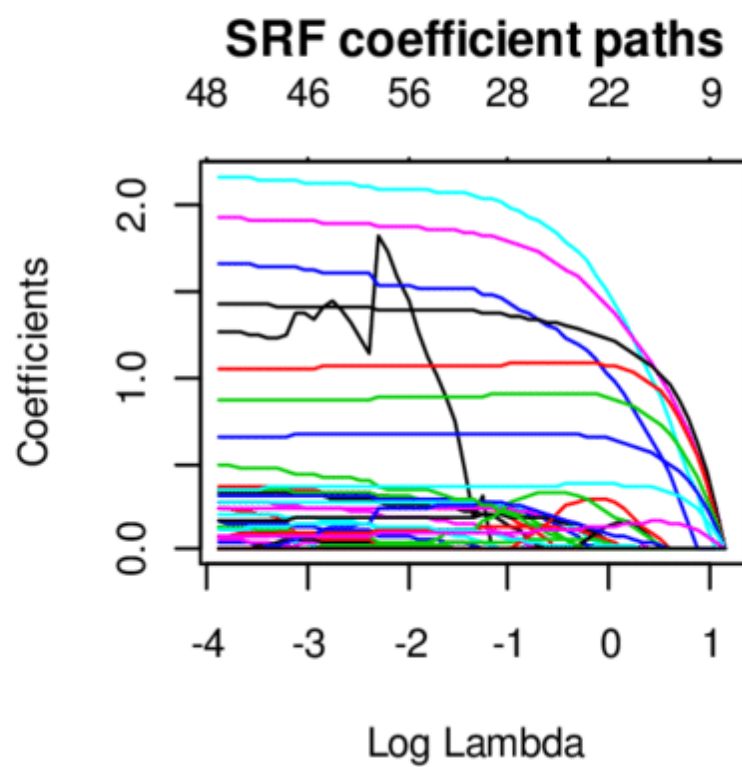
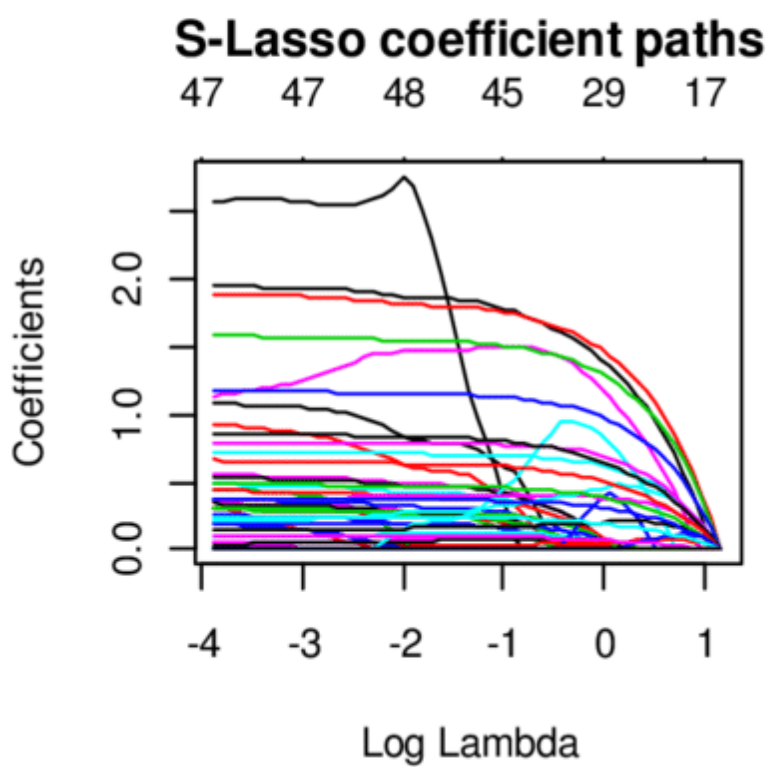
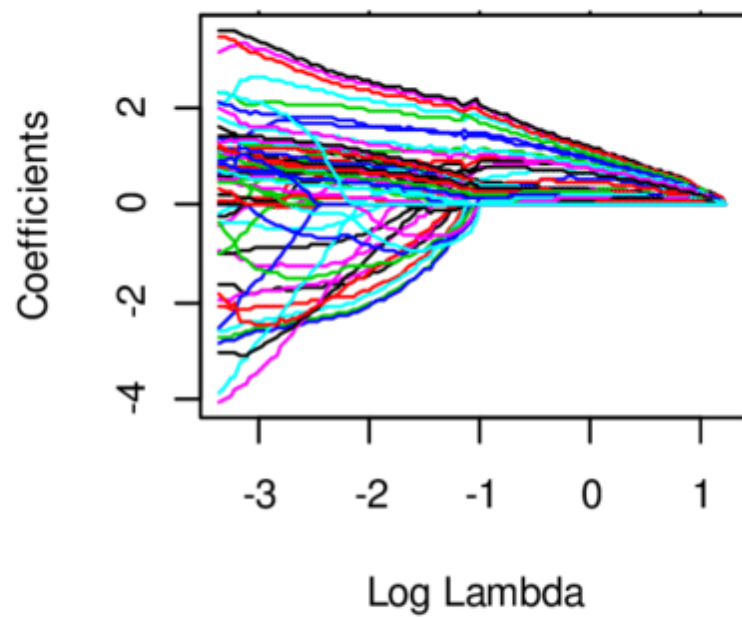
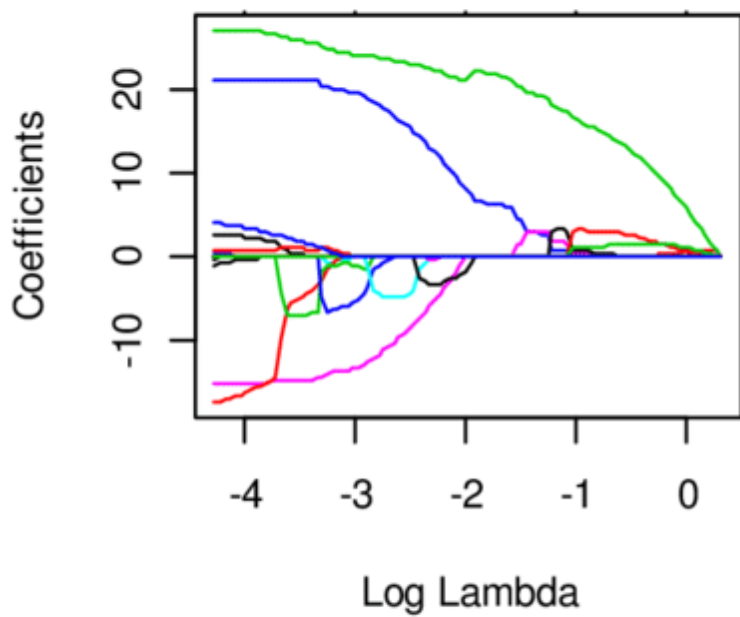
Multicollinear regression data is often evaluated using the principle components regression approach. The significant components regression approach, like ridge regression, reduces standard errors by biasing the regression estimates. Principal component analysis (PCA) is used first to modify the training data, and then the resulting transformed samples are used to train the regressors.

9. Partial Least Squares Regression

The partial least squares regression technique is a fast and efficient covariance-based regression analysis technique. It is advantageous for regression problems with many independent variables with a high probability of multicollinearity between the variables. The method decreases the number of variables to a manageable number of predictors, then is utilized in a regression.

10. Elastic Net Regression

Elastic net regression combines ridge and lasso regression techniques that are particularly useful when dealing with strongly correlated data. It regularizes regression models by utilizing the penalties associated with the ridge and lasso regression methods.



[IMAGE](#)

Frequently Asked Questions

Q1. What is a regression model with an example?

A. A regression model is a statistical tool used to examine the relationship between a dependent variable and one or more independent variables. For example, consider a real estate dataset where the dependent variable is the house price, and the independent variables are features like the size, number of bedrooms, and location. Using regression analysis, we can create a model that predicts house prices based on these features. The model estimates the coefficients for each independent variable, allowing us to understand how much each feature contributes to the final house price prediction, aiding in pricing decisions and market analysis.

Q2. Why do we use regression model?

A. We use regression models to analyze the relationship between a dependent variable and one or more independent variables. These models help us understand how changes in the independent variables impact the dependent variable. Regression is widely used for prediction, hypothesis testing, and understanding the influence of factors, making it a valuable tool in various fields such as economics, finance, and social sciences.