

Lasso Regression

The lasso regression model was originally developed in 1989.

The LASSO (Least Absolute Shrinkage and Selection Operator) is a regression method that involves penalizing the absolute size of the regression coefficients. It adds a factor of sum of absolute value of coefficients in the optimization objective.

It is a type of linear regression that uses shrinkage. Shrinkage is where data values are shrunk towards a central point, like the mean. This particular type of regression is for models showing high levels of multicollinearity or when you want to automate certain parts of model selection, like variable selection/parameter elimination.

Lasso regression performs L1 regularization, which adds a penalty equal to the absolute value of the magnitude of coefficients. This type of regularization can result in sparse models with few coefficients. Some coefficients can become zero and eliminated from the model. Larger penalties result in coefficient values closer to zero, which is the ideal for producing simpler models. This makes the Lasso far easier to interpret than the Ridge.

$$\sum_{i=1}^n (y_i - \sum_j x_{ij} \beta_j)^2 + \lambda \sum_{j=1}^p |\beta_j|$$

A tuning parameter, λ controls the strength of the L1 penalty. λ is basically the amount of shrinkage:

- When $\lambda = 0$, no parameters are eliminated. The estimate is equal to the one found with linear regression.
- As λ increases, more and more coefficients are set to zero and eliminated (theoretically, when $\lambda = \infty$, *all* coefficients are eliminated).
- As λ increases, more coefficients are reduced to zero that is fewer predictors is selected and there is more shrinkage of the non-zero coefficient. Bias increases and variance decreases as lambda increases.
- As λ decreases, variance increases.

Lasso Regression is helps in shrinking the regression coefficient and can reduce variance without a substantial increase in bias.

Second, Lasso Regression can increase model interpretability.

Often some of the explanatory variables in an OLS multiple regression analysis are not really associated with the response variable. As a result, model is over fitted and more difficult to interpret. With Lasso Regression, the regression coefficients for unimportant variables are reduced to zero, which effectively removes them from the model and produces a simpler model that selects only the most important predictors.