The Lasso

Same as Ridge Regression which penalize linear regression, but the main disadvantage of the ridge regression is it will shrink the coefficients but not set any of them to zero which can be a challenge when **inference and interpretation** is needed or selecting the features.

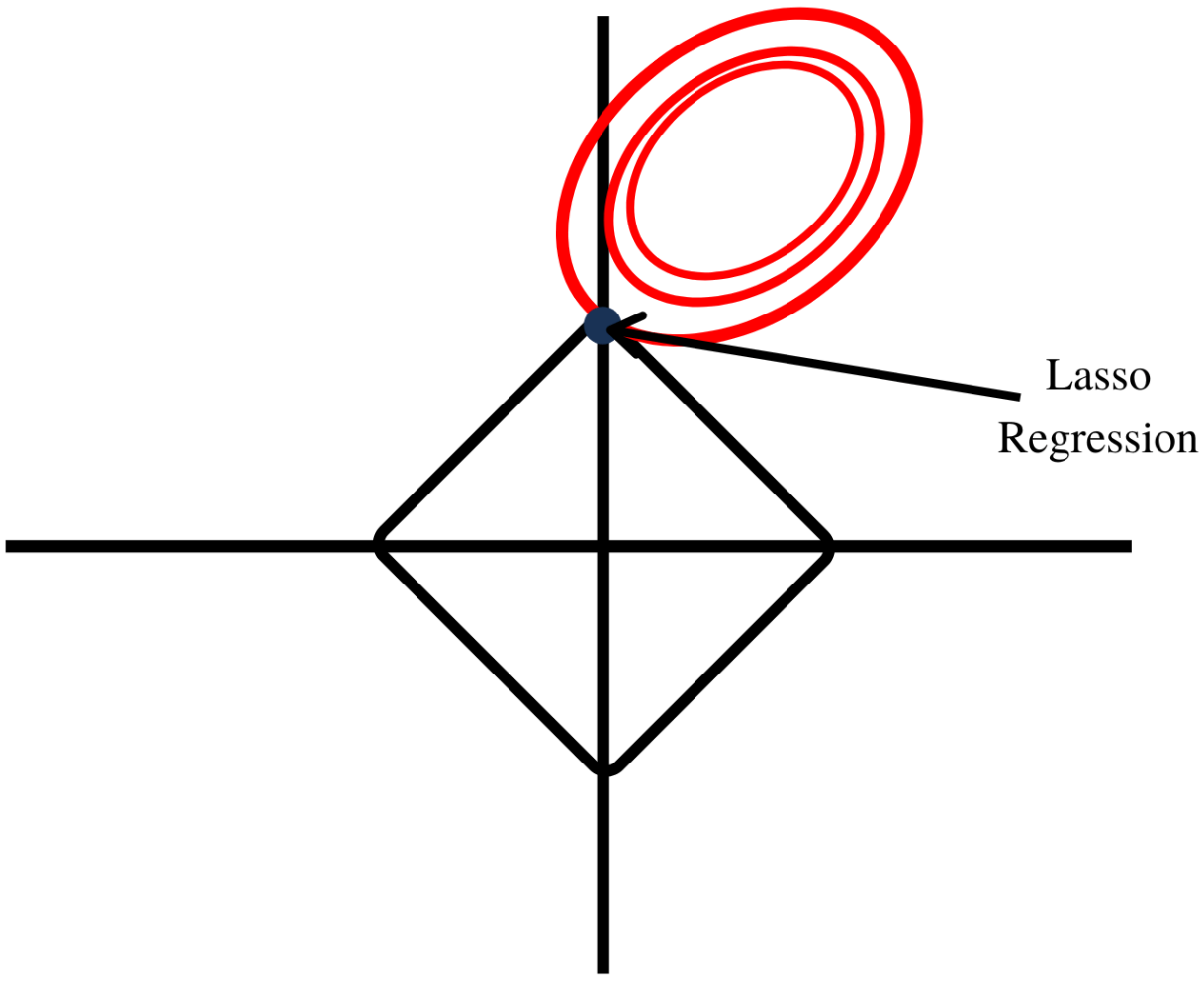
The **Ridge Regression** main motivation behind was to deal with :

- High Multicollinearity
- High Dimensionality
- Prediction Accuracy

And it used the **Squared Euclidean Norm** which is the L_2 Norm, they used it for an arbitrary reason behind which lead for a consideration in other Norms such as L_1 which is called **The Lasso Regression**

Lasso Vs Ridge

- The Ridge Regression uses L_2 Norm
- The Lasso Regression uses L_1 Norm



-

Lasso Regression

It's introduce a penalty term same as the Ridge Regression but in the L_1 Norm which uses :

$$f_{pen}(\beta, \lambda) = \lambda_1 \|\beta\|_1$$

Which give us the **Lasso Cost Function**

$$\mathcal{L}_{\text{lasso}}(\beta; \lambda) = \|Y - X\beta\|_2^2 + \lambda_1 \|\beta\|_1 = \sum_{i_1}^{n} (Y_i - X_i\beta)^2 + \lambda_1 \sum_{j=1}^{p} |\beta_j|$$

- Contains the **Least Squares** and **Regularization Term**
- Absolute value doesn't have a solution at 0 so no close-form solution exist unlike **Ridge Regression**
-