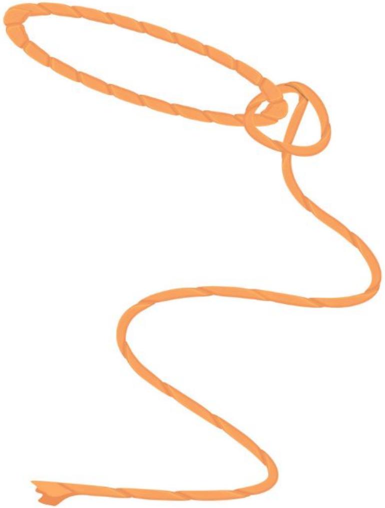
# Lasso/Ridge Regression



Salma Tahlil, Susan Lu,
Mason Lonoff, Rob Dewan

# Lasso Regression

**Lasso** is an acronym for least absolute shrinkage and selection operator

- Lasso regression is a Linear Regression model but with a penalty which is also called regularization
- Lasso uses the shrinkage technique, in this technique the data values are shrunk towards a central point such as the mean value

$$\text{minimize} \quad \underbrace{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}_{\text{RSS}} + \lambda \underbrace{\sum_{j=1}^{p}|\hat{\beta}_j|}_{\text{penalty term}}$$

$p$ = # coefficients for predictors
$\lambda$ is a tuning parameter $\geq 0$

# Ridge Regression

Ridge regression is a modeling method that minimizes the following:

$$\mathrm{RSS} + \lambda \sum_{j=1}^{p} \beta_j^2$$

RSS is the residual square sum, which is what a linear regression would minimize.

$\beta_j$ represents the correlation of each variable. p symbolizes how many variables there are. $\lambda$ is the amount the user can desensitize the slopes. The higher $\lambda$ is set to be, the less reliant on the slope the data would be. If $\lambda$ is too high, the slopes would ultimately be reduced down close to 0.

# Advantages

Usually results in lower variability when testing the model against testing set by residing the sensitivity of the slope and this helps reduce overfitting

Can also be used to create models when there are more variables than data points using cross validation

The main advantage of Lasso is the automatic feature selection, helps decide which features should and should not be included on its own

Both regression models are well-suited for models that show a high level of multicollinearity

# Disadvantages

Raises bias for training data set

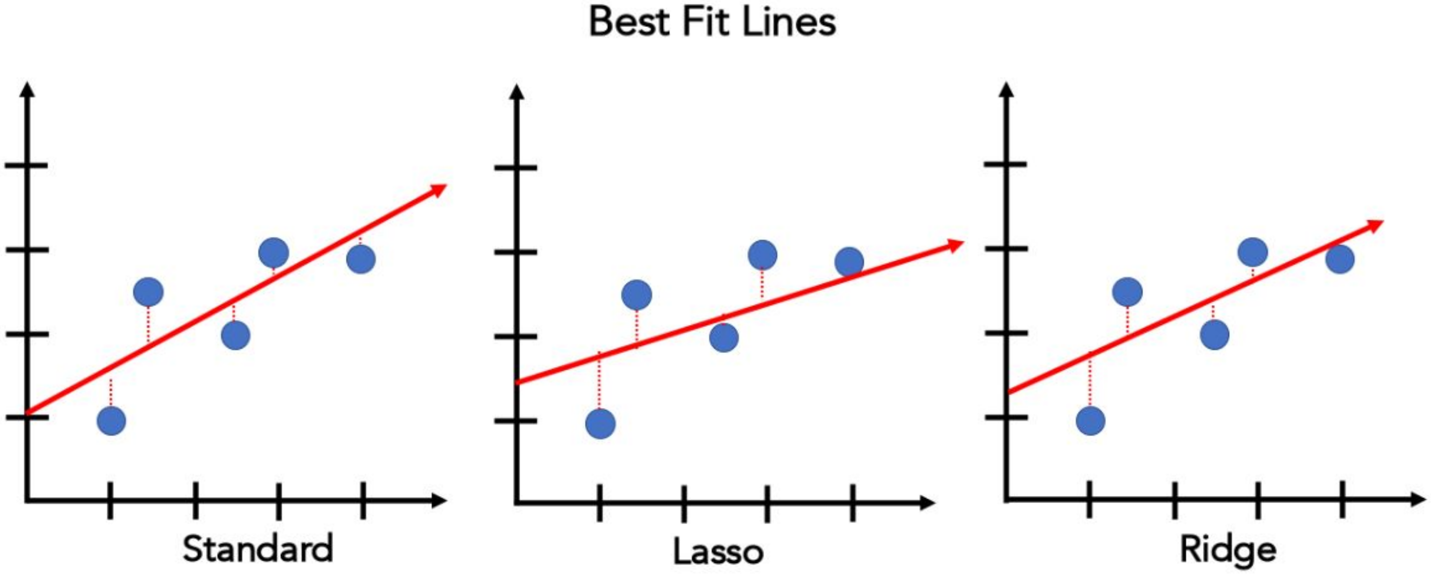
Requires the user to choose the weight of the penalty term

Difficult to estimate standard error

Ridge regression can be hard to interpret as there’s no feature selection

Lasso regression may underperform Ridge regression when there are many correlated predictors in the model

# Differences between Lasso/Ridge and Linear Regression



# Data processing steps

- Standardization is necessary
  - scaled_feature = (feature - mean(feature)) / std(feature)
  - Also seen used:
    - ridge = Ridge(normalize=True)
    - lasso = Lasso(normalize=True)
- Every column must contain numerical data
  - For categorical data, create dummy variables
- All missing values must be imputed
- Both Ridge and Lasso Regression can tolerate outliers better than Linear Regression, so it is not necessary to remove them

# What hyperparameters can be tuned? What does each one represent?

The main hyperparameter in Lasso and Ridge Regression is lambda $\lambda$

**Lamda** controls how large the penalty is, by increasing lambda, we increase the constraint on the size of the beta vector

- A larger value chosen for lambda will result in a greater quantity of bias into the algorithms output
- Lambda is known as “alpha” when using python

Scikit learn provides built in methods for hyperparameter tuning:

- **Grid search** (GridSearchCV): this function will test each hyperparameter combination against the model in order to determine the combination with the best score value
- **Randomized search** (RandomizedSearchCV): takes continuous random variable object as its param_grid argument

Other hyperparameters you may come across:

- max_iter, normalize, cv (cross-validation), fit_intercept,

# Appendix

StatQuest explanation of Ridge Regression:

https://www.youtube.com/watch?v=Q81RR3yKn30

Lasso and Ridge Regression Tutorial | DataCamp

Lasso and Ridge regression: An intuitive comparison | by Thomas Le Menestrel | Towards Data Science

Ridge Regression From Scratch In Python [Machine Learning Tutorial] - YouTube

"Lasso & Ridge Regression" in 200 Words - Data Science (thaddeus-segura.com)

StatQuest comparing Ridge and Lasso Regressions:

https://www.youtube.com/watch?v=Xm2C_gTAl8c

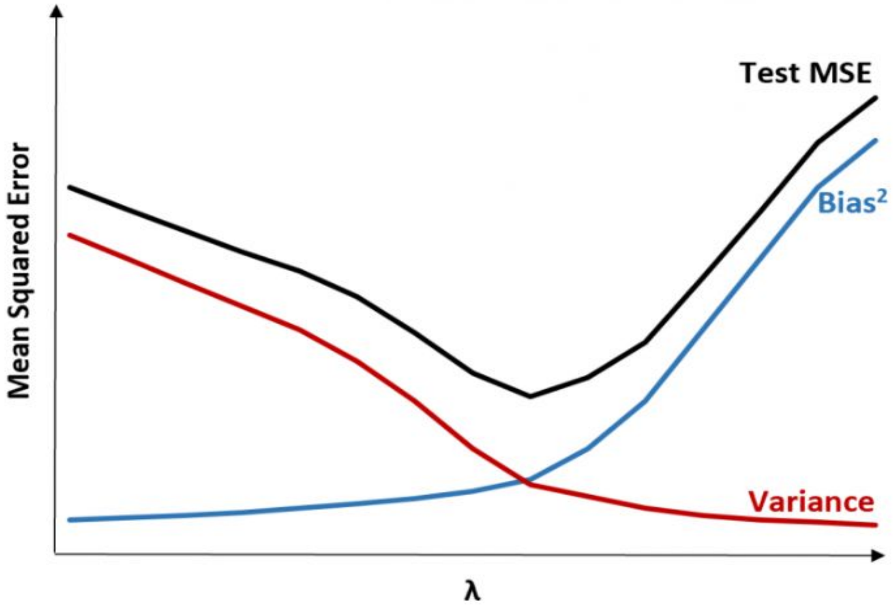
Lasso Regression for Beginners | By Dr. Ry @Stemplicity - YouTube

Ridge Regression for Beginners! - YouTube

http://www.science.smith.edu/~jcrouser/SDS293/labs/lab10-py.html#6.6.2-The-Lasso

# Appendix Continued

Taken from: [Introduction to Ridge Regression - Statology]



# Appendix Continued

Visualization of how weights are affected as alpha (aka lambda) is increased.