Ridge & Lasso Regression

Ridge Regression

→ Cost function in case of Ridge Regression.

$$CF = \underbrace{\sum_{i=1}^{n} \left(y_i - \sum_j x_{ij} \beta_j\right)^2}_{\text{From Linear Regression.}} + \underbrace{\lambda \sum_{j=1}^{p} \beta_j^2}_{L_2 \text{ Regularization.}}$$

→ This reduces variance in Linear Regression Model.
→ This increases bias in Linear Regression Model.
→ This prevents overfitting.
→ λ learned through cross-validation / hyperparameter tuning.
→ This reduces model complexity. & over-fitting.

→ This puts contraints on the coefficients (w)
→ λ (penalty term) reduces overall slope
→ So, Ridge Regression shrinks coefficients and helps to reduce model complexity & multi-colinearity.

Lasso Regression

→ Cost function in case of Lasso Regression.

$$CF = \underbrace{\sum_{i=1}^{n} \left(y_i - \sum_j x_{ij} \beta_j\right)^2}_{\text{From Linear Regression}} + \underbrace{\lambda \sum_{j=1}^{p} |\beta_j|}_{L_1 \text{ Regularization.}}$$

→ Reduces variance, Increases bias, Reduces overfitting. & complexity
→ λ (penalty term) reduces overall slope and is learnt through cross-validation / hyperparameter tuning.
→ constraints β
→ Reduces multi-collinearity

→ Helps in feature selection as reduces slope of some (coeffi-
cients of some) features to 0 (zero).