Ridge vs. Lasso Regression

Ridge

+ Reduces Multicollinearity
+ Continuous Shrinking
+ Stable Solutions
+ Computationally Efficient

- No variable selection
- Interpretability
- Sensitive to scale

Lasso

+ Variable selection
+ Sparse models
+ Improves interpretability
+ Particularly useful for when $p > n$

- Collinearity issues
- Bias in coefficients (ℓ_1 penalty is harsher)
- Computationally intensive

λ Tuning

- *K*-fold Cross Validation
    1. Choose the number of folds *K*
    2. Split the data accordingly into training and testing sets.
    3. Define a grid of values for λ
    4. For each λ, calculate the validation MSE within each fold
    5. For each λ, calculate the overall cross-validation MSE
    6. Locate under which λ cross-validation MSE is minimized, i.e. minimum_cv λ
- Packages such as will `glmnet` do this automatically