Summary

This week's lecture covered methods for improving linear models through model selection and regularization. We examined techniques like subset selection, ridge regression, and the lasso, all of which aim to either improve prediction accuracy or make the model more interpretable by reducing complexity.

Concepts

- **Subset selection**: Involves choosing a subset of predictors that best explain the response. Includes best subset, forward stepwise, and backward stepwise selection.

- **Best subset selection**: Fits all possible models and selects the best one for each model size. Computationally expensive for large ppp.

- **Forward stepwise selection**: Starts with no variables and adds one at a time based on model improvement.

- **Shrinkage methods**: Ridge regression and lasso are techniques that shrink coefficient estimates toward zero to reduce variance.

- **Ridge regression**: Penalizes the sum of squared coefficients (L2 norm). Useful when predictors are highly correlated. All coefficients remain nonzero.

- **Lasso**: Uses an L1 penalty which can set some coefficients exactly to zero, helping with variable selection.

- **Bias-variance trade-off**: Regularization helps reduce variance at the cost of adding some bias, leading to better generalization.

Uncertainties

One area I'm still unsure about is how to decide between using ridge regression and lasso in practice, especially when the goal is both accurate prediction and variable selection. I also found it a bit tricky to understand how the different evaluation criteria compare and when one is preferred over another.