CHE384, From Data to Decisions: Measurement,
Uncertainty, Analysis, and Modeling

Lecture 51
Addressing Multicollinearity

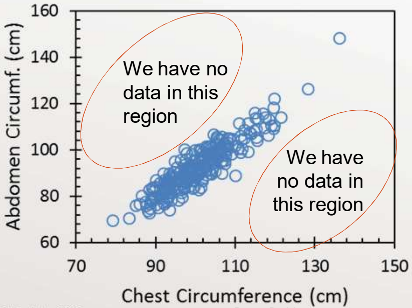
Chris A. Mack
Adjunct Associate Professor

http://www.lithoguru.com/scientist/statistics/

What to Do About Multicollinearity

- If we only care about prediction, restrict the scope of the model to coincide with the range of predictor variables that exhibit the same pattern of multicollinearity
- Drop some correlated predictor variables (ones with the highest VIF)
- Add data cases that break the pattern of multicollinearity
- Measure some coefficients in a separate experiment (then fix those coefficients)
- Use Ridge Regression
- Use Principal Component Analysis (PCA)

Adding Data to Break Collinearity



Two possibilities:
1. We have no data because we did not collect it
2. We have no data because they do not exist

What to Do About Multicollinearity

- Use Ridge Regression
  - Add a small bias to the estimators to help break the correlations and reduce variance
  - Find the minimum of

penalty

$$S=\sum_{i=1}^{n}\left(y_i-\beta_0-\sum_{j=1}^{p}\beta_j x_{ji}\right)^2+k\sum_{j=1}^{p}(\beta_j)^2$$

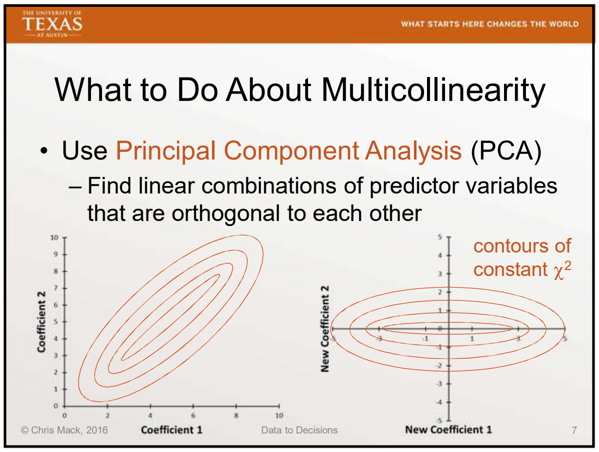
  - The penalty shrinks our coefficients
  - We generally scale and center (standardize) our data before we perform the regression

Ridge Regression

- To preform ridge regression, just add p–1 rows to the standardized data, with 0 for the y-values and $\sqrt{k}\boldsymbol{I}_p$ for the x-values (predictor values), then do a standard OLS regression
- How to pick k, the ridge coefficient?
  - Plot the fit coefficients versus k (called a *ridge trace*), pick the smallest k at which the coefficients start to level off (don't change much)
  - Using a validation data set, find k that minimizes validation SSE

Properties of Ridge Regression

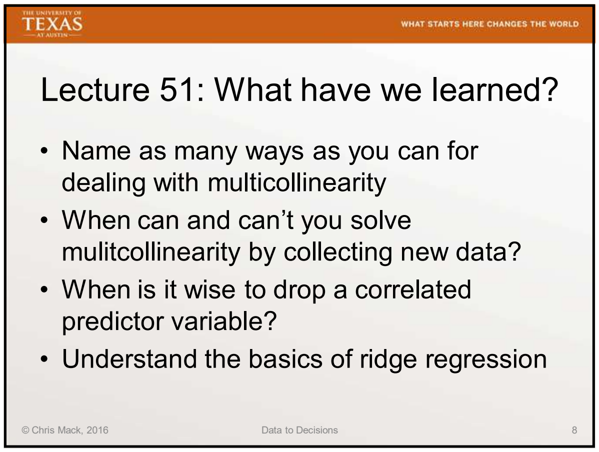
- The ridge regression model coefficients are biased smaller, but have smaller variance
  - As $k \to 0$, $\hat{\beta}_{ridge} \to \hat{\beta}_{OLS}$
  - As $k \to \infty$, $\hat{\beta}_{ridge} \to 0$
  - For orthonormal design matrix, $\hat{\beta}_{ridge} = \hat{\beta}_{OLS}/(1+k)$
- Even for perfect multicollinearity, the ridge regression solution will always exist
- There is always a value of k where the MSE of the ridge coefficients are smaller than for OLS



What to Do About Multicollinearity

- Use Principal Component Analysis (PCA)
  - Find linear combinations of predictor variables that are orthogonal to each other

contours of constant χ^2



Lecture 51: What have we learned?

- Name as many ways as you can for dealing with multicollinearity
- When can and can't you solve mulitcollinearity by collecting new data?
- When is it wise to drop a correlated predictor variable?
- Understand the basics of ridge regression