Section for Applied Statistics and Data Analysis

TA: Cong Mu

Office Hour: Wednesday 10:00AM - 12:00PM

November 22, 2019

Overview

1. Some Statistics
   - Changes of Scale
   - Collinearity
   - Weighted Least Squares
2. Some Programming
   - Examples in Faraway

Changes of Scale

- A change of scale is often helpful when values of variables are all very large or all very small
  - Interpretation
  - Numerical stability
- $x_i \to \frac{x_i + a}{b}$
  - t-statistic, F-statistic, R^2 and $\widehat{\sigma}^2$ unchanged
  - $\widehat{\beta}_i \to b\widehat{\beta}_i$
- $y \to \frac{y + a}{b}$
  - t-statistic, F-statistic and R^2 unchanged
  - $\widehat{\beta} \to b\widehat{\beta}$, $\widehat{\sigma} \to b\widehat{\sigma}$

Collinearity

- Collinearity leads to imprecise estimates of β
- Detection of collinearity
  - Correlation matrix of the predictors
  - Regression of x_i on all other predictors
  - Eigenvalues of $X^\top X$
  - Variance inflation factor $\frac{1}{1-R_i^2}$

Weighted Least Squares

- $Var[\epsilon] = \sigma^2 I \rightarrow Var[\epsilon] = \sigma^2 \Sigma$
- **Generalized Least Squares**
  The errors are dependent

$$\widehat{\beta} = \left(X^\top \Sigma^{-1} X\right)^{-1} X^\top \Sigma^{-1} y, \quad Var\left[\widehat{\beta}\right] = \sigma^2 \left(X^\top \Sigma^{-1} X\right)^{-1}.$$

- **Weighted Least Squares**
  The errors are independent but not identically distributed

$$\Sigma = \text{diag}\left(\frac{1}{w_1}, \cdots, \frac{1}{w_n}\right).$$

Examples in Faraway Chapter 7 & 8

- **Example:** `savings` dataset
- **Example:** `seatpos` dataset
- **Example:** `fpe` dataset

Thanks for listening!