Theory of Linear Models: HW 3

Fall 2024

Due: Thursday, 11/21 at 11:59pm

1. **F-test** Assume $Y = X\beta + \varepsilon$, $\varepsilon \sim \mathcal{N}(0, \sigma^2 I_n)$. The hypothesis of interest is $H_0 : A\beta = c$ for a fixed matrix A and vector c. Recall that under the null hypothesis, the LSE is $\hat{\beta}_H = \beta + (X^T X)^{-1} A^T [A(X^T X)^{-1} A^T]^{-1} (c - A\hat{\beta})$

   a. Find B such that $RSS_H - RSS = Z^T B^{-1} Z$. for $Z = A\hat{\beta} - c$.
   b. Verify that $RSS/\sigma^2 \sim \chi^2_{n-p}$.
   c. Under H_0, what is the distribution of $\frac{RSS_H - RSS}{\sigma^2}$?
   d. Prove that RSS and RSS_H are independent.
   e. What is the distribution of $\frac{(RSS_H - RSS)/q}{Rss/(n-p)}$?
   f. Show that when $c = 0$,

$$\frac{(RSS_H - RSS)/q}{Rss/(n-p)} = \frac{Y^T (P_X - P_H) Y/q}{Y^T (I - P_X) Y/(n-p)},$$

2. **Adding covariates** Let $R = I - P_X, L = (X^T X)^{-1} X^T Z, M = (Z^T R Z)^{-1}$. This problem is regarding the theorem we discussed on adding covariates.

   a. Part 1 of our theorem states $\hat{\gamma} = (Z^T R Z)^{-1} Z^T R Y$. Provide a full proof of this result, filling in the missing steps from class. Make sure to include finding the LSE estimate of δ in the model $Y = W\delta + \varepsilon$, for W the concatenated design matrix $W = (X, Z)$.

   Hint: Find an expression for $W^T W$ in terms of X, R, and Z.

   b. Prove the formulas given for $\mathrm{Var}(\hat{\gamma})$ and $\mathrm{Cov}(\hat{\gamma}, \beta_{new})$.

   c. Show $\hat{\beta}_{new} = \hat{\beta}_{old} - L\hat{\gamma}$.

3. **Removing an observation.** Verify the formula for updating $\hat{\beta}$ when removing observation i:

$$\hat{\beta}_{(i)} = \hat{\beta} - \frac{(X^T X)^{-1} x_i (Y_i - X_i^T \hat{\beta})}{1 - h_{ii}}.$$

4. **Underfitting** Suppose the true data generating model is $Y = X\beta + Z\gamma + \varepsilon$, and we instead fit the model $Y = X\beta + \varepsilon$ to compute $\hat{\beta}$.

   a. What is the bias of $\hat{\beta}$ as an estimator of the true β?
   b. What is the variance of $\hat{\beta}$?
   c. Let $\hat{\beta}_C$ represent the estimate of β from the true data generating model. What is $\mathrm{Var}(\hat{\beta}_C)$?
   d. Show that the inclusion of the Z covariates reduces the variance of the estimator for β. I.e., show that $\mathrm{Var}(\hat{\beta}_C) < \mathrm{Var}(\hat{\beta})$.

5. **Ridge Metric** Consider observed data (X, Y) of dimensions $n \times p$ and $n \times 1$ respectively. Assume X has singular value decomposition $X = UDV^T$, where U and V are orthogonal matrices of dimensions $n \times p$ and $p \times p$ respectively, and D is a $p \times p$ diagonal matrix.

a. What is the distribution of $Z = U^T Y$ when $Y \sim \mathcal{N}(0, \Sigma)$?

b. Find an alternative expression for the ridge weight matrix $W_\lambda = (X^T X + \lambda I_p)^{-1}$ using the Sherman-Morrison-Woodbury formula and the singular value decomposition of X.

c. Find an expression for the statistic $T_\lambda = Y^T X W_\lambda X^T Y$ in terms of Z and the singular values $d_i, i = 1, \ldots, p$.

d. Assuming $Y \sim \mathcal{N}(0, \sigma^2 I_n)$, what is the distribution of T_λ?