Homework 6

1. Consider the linear model $y_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2)$ for $i = 1, 2, 3, 4$. Before you collect the four data points, you have freedom to choose x_1, x_2, x_3, x_4 and then observe y_1, y_2, y_3, y_4. Suppose you can only choose x_1, x_2, x_3, x_4 in the range $[-1, 1]$. How would you choose x_1, x_2, x_3, x_4 so that the accuracy of the MLE/LSE of β_1 is the best?

2. We have derived the formulas for $\hat{\beta}_0$ and $\hat{\beta}_1$ by setting the derivatives to zeros and solving the linear equations. Show that $\hat{\beta} = (X^T X)^{-1} X^T y$ leads to the same formula when $p = 2$.

3. We have derived the means, variances and covariance of $\hat{\beta}_0$ and $\hat{\beta}_1$ using straightforward calculations. Show that $\hat{\beta} \sim N(\beta, \sigma^2 (X^T X)^{-1})$ leads to the same formulas.

4. Consider $y \sim N(X\beta, \sigma^2 I_n)$. Use $\sigma_c^2 = c\|y - X\hat{\beta}\|^2$ to estimate σ^2, where $\hat{\beta}$ is the LSE. We know that $\frac{\|y - X\hat{\beta}\|^2}{\sigma^2} \sim \chi^2_{n-p}$.

   (a) What c leads to an unbiased estimator?

   (b) What c gives you the MLE?

   (c) Find the c that minimizes $\mathbb{E}(\hat{\sigma}_c^2 - \sigma^2)^2$.

   (d) Use the fact $\frac{\|y - X\hat{\beta}\|^2}{\sigma^2} \sim \chi^2_{n-p}$ to build a confidence interval of σ^2.

5. Consider $y \sim N(X\beta, \sigma^2 I_n)$. For a new data point $x^* = (1, x_1^*, x_2^*, ..., x_{p-1}^*)^T$, we can predict its response by $(x^*)^T \hat{\beta}$, where $\hat{\beta}$ is the MLE.

   (a) Find the distribution of $(x^*)^T \hat{\beta}$.

   (b) Construct a confidence interval of $(x^*)^T \beta$ using t-distribution.

6. *(Pythagorean Identity.)* Review and prove the following lists of Pythagorean identities:

   (a) For any random variable $\hat{\theta}$ and any number θ, $\mathbb{E}(\hat{\theta} - \theta)^2 = \text{var}(\hat{\theta}) + (\mathbb{E}\hat{\theta} - \theta)^2$.

   (b) For any real numbers $x_1, ..., x_n, \theta$, $\sum_{i=1}^n (x_i - \theta)^2 = \sum_{i=1}^n (x_i - \bar{x})^2 + n(\bar{x} - \theta)^2$.

   (c) For any vectors $X_1, ..., X_n, \theta$, $\sum_{i=1}^n \|X_i - \theta\|^2 = \sum_{i=1}^n \|X_i - \bar{X}\|^2 + n\|\bar{X} - \theta\|^2$.

   (d) For any full rank $X \in \mathbb{R}^{n \times p}$, $y \in \mathbb{R}^{n \times 1}$ and $\mathbb{R}^{p \times 1}$, $\|y - X\beta\|^2 = \|y - X\hat{\beta}\|^2 + \|X\hat{\beta} - X\beta\|^2$, where $\hat{\beta} = (X^T X)^{-1} X^T y$.

   (e) Can you explain the common phenomenon?