Tutorial & Practical 1: Linear Regression

Question 1

Given the model

$$\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where $\mathbf{y} \in \mathbb{R}^n$, $X \in \mathbb{R}^{n \times p}$ is full rank p and $\boldsymbol{\epsilon} \in \mathbb{R}^n \sim \mathcal{N}(0, \sigma^2 I_n)$. Let $\hat{\boldsymbol{\beta}}$ be the estimate of $\boldsymbol{\beta}$ obtained by least square estimation and $H = X\left(X^\top X\right)^{-1} X^\top$ the orthogonal projector onto the subspace spanned the columns of X

- Write $U = \mathbf{y} - X\hat{\boldsymbol{\beta}}$ as a function of H and $\boldsymbol{\epsilon}$
- Write $V = X\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)$ as a function of H and $\boldsymbol{\epsilon}$
- Using the results obtained find $\mathrm{Cov}\left(U, V\right)$
- Use this results to show that

$$\left(\mathbf{y} - X\boldsymbol{\beta}\right)^\top \left(\mathbf{y} - X\boldsymbol{\beta}\right) = \left(\mathbf{y} - X\hat{\boldsymbol{\beta}}\right)^\top \left(\mathbf{y} - X\hat{\boldsymbol{\beta}}\right) + \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)^\top X^\top X \left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right)$$

- What is the conclusion you can make from the left side of the above equation?
- Using the above decomposition find the distribution of $\left(\mathbf{y} - X\hat{\boldsymbol{\beta}}\right)^\top \left(\mathbf{y} - X\hat{\boldsymbol{\beta}}\right)/\sigma^2$

Question 2

Let

$$\mathbf{y} = \mathbf{1}_n \beta + \boldsymbol{\epsilon}$$

where $\mathbf{y} \in \mathbb{R}^n$ and $\boldsymbol{\epsilon} \in \mathbb{R}^n \sim \mathcal{N}(0, \sigma^2 I_n)$.

- Find of the least square estimate of β?
- What does represent $\hat{\beta}$ and the associated $\hat{\mathbf{y}}$
- Provide the expression of the projector generating $\hat{\mathbf{y}}$
- Use this projector to provide an alternative expression for $\sum_{i=1}^n \left(y_i - \bar{y}\right)^2$
- Show that $\bar{y}$ is independent of $\sum_{i=1}^n \left(y_i - \bar{y}\right)^2$
- Show that $\sum_{i=1}^n \left(y_i - \bar{y}\right)^2 / \sigma^2$ is χ^2_{n-1}

Question 3

Let

$$\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where $\mathbf{y} \in \mathbb{R}^n$, $X \in \mathbb{R}^{n\times p}$ is full rank p and $\boldsymbol{\epsilon} \in \mathbb{R}^n \sim \mathcal{N}(0, \sigma^2 I_n)$.

- Give the expression of the log-likelihood $\ell\left(\boldsymbol{\beta}, \sigma^2\right)$
- Show that the least square estimate for $\boldsymbol{\beta}$ is also the maximum likelihood estimator for $\boldsymbol{\beta}$
- Find the maximum likelihood estimator of σ^2
- Using the expression of $\hat{\boldsymbol{\beta}}$ and $\hat{\sigma}^2$ provide the expression of $\ell\left(\hat{\boldsymbol{\beta}}, \hat{\sigma}^2\right)$
- Taking $\boldsymbol{\theta} = \left(\boldsymbol{\beta}^\top, \sigma^2\right)^\top$ derive the expression of the information matrix $\mathbf{I} = -\mathbb{E}\left[\frac{\partial^2 \ell}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top}\right]$
- Discuss the efficiency of $\hat{\boldsymbol{\beta}}$ and $\hat{\sigma}^2$

Question 4

Let

$$\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where $\mathbf{y} \in \mathbb{R}^n$, $X \in \mathbb{R}^{n\times p}$ is full rank p and $\boldsymbol{\epsilon} \in \mathbb{R}^n \sim \mathcal{N}(0, \sigma^2 I_n)$.
Furthermore to the above model there are a number of restrictions or constraints that $\boldsymbol{\beta}$ need to satisfies, these constraints are described by

$$A\boldsymbol{\beta} = \mathbf{c}$$

where A is a known $q \times p$ matrix of rank q and $\mathbf{c}$ is a known $q \times 1$.

Find an estimate of $\boldsymbol{\beta}$ that minimizes the squared error subject to these constraint (use Lagrange multipliers, one for each constraint $\mathbf{a}_i\boldsymbol{\beta} = c_i$, $i = 1, ..., q$).

Question 5

Let

$$\mathbf{y} = X\boldsymbol{\beta} + \boldsymbol{\epsilon}$$

where $\mathbf{y} \in \mathbb{R}^n$, $X \in \mathbb{R}^{n\times p}$ is full rank $r < p < n$ and $\boldsymbol{\epsilon} \in \mathbb{R}^n \sim \mathcal{N}(0, \sigma^2 I_n)$.

- Write the least square estimate of $\boldsymbol{\beta}$ using the $n \times n$ matrix S, the diagonal matrix $\Sigma = \text{diag}\left(\sigma_1, ..., \sigma_r\right)$ and the $p \times p$ matrix Q, obtained from the singular value decomposition of X, $X = S\Sigma Q^\top$
- Write the least square estimate of $\boldsymbol{\beta}$ as a function σ_i, $\mathbf{s}_i$ and $\mathbf{q}_i$ where $\mathbf{s}_i$ and $\mathbf{q}_i$ are the column vectors of S and Q respectively.

- Consider the estimate of $\boldsymbol{\beta}$ which can be represented as a projection $\boldsymbol{\beta}_p = Q_k z$ where $Q_k = (\mathbf{q}_1, ..., \mathbf{q}_k)$, what is the relation between the least square estimate of $\hat{\boldsymbol{\beta}}_{LS}$ and $\boldsymbol{\beta}_p$ where z is obtained by minimizing $\|\mathbf{y} - XQ_k z\|^2$
- Provide an expression for the residual $R_k = \|\mathbf{y} - X\hat{\boldsymbol{\beta}}_p\|^2$