Tutorial Notes 4 of MATH3424

1 Summary of course material

1.1 Multiple Linear Regression

- Model:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p + \epsilon$$

Suppose there are n observations, each of them can be written as

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_p x_{ip} + \epsilon_i, \quad i = 1, \cdots, n$$

- Matrix notation:

$$\mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}_{n\times 1}, \quad \mathbf{X} = \begin{pmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}_{n\times(p+1)}, \quad \boldsymbol{\beta} = \begin{pmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{pmatrix}_{(p+1)\times 1}, \quad \boldsymbol{\epsilon} = \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_n \end{pmatrix}_{n\times 1}$$

Then the observed data can written as

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

$S(\beta) = \|y - X\beta\|^2$

- Parameter estimation (Least square estimator):

$\nabla_\beta S(\beta) = X^T \cdot 2(X\beta - y) = 0$

$\Rightarrow X^T X\beta = X^T y \Rightarrow \beta = (X^T X)^{-1} X^T y$

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{y}$$

If $\epsilon_1, \cdots, \epsilon_n$ are i.i.d. with common variance σ^2, an unbiased estimate of σ^2 is given by

$$\hat{\sigma}^2 = \frac{\text{SSE}}{n - \underline{(p+1)}} = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - p - 1}$$

- Centering and Scaling: (of predictors).

Centering does not affect the regression coefficients except that the estimate of constant term $\hat{\beta}'_0$ is always 0.

Scaling will change the values of the regression coefficient.

2 Questions

1. Using the following summary statistics:

$n = 20,\quad \sum_{i=1}^{20} x_{i1} = 114,\quad \sum_{i=1}^{20} x_{i2} = -136,\quad \sum_{i=1}^{20} y_i = 222,$

$\sum_{i=1}^{20} x_{i1}^2 = 860,\quad \sum_{i=1}^{20} x_{i1}x_{i2} = -1025,\quad \sum_{i=1}^{20} x_{i2}^2 = 1228,\quad \sum_{i=1}^{20} x_{i1}y_i = 1537,$

$\sum_{i=1}^{20} x_{i2}y_i = -1824,\quad \sum_{i=1}^{20} y_i^2 = 2950,$

$S_{x_1x_1} = 210.2\quad S_{x_1x_2} = -249.8\quad S_{x_2x_2} = 303.2\quad S_{x_1y} = 271.6$

$S_{x_2y} = -314.4\quad S_{yy} = 485.8$

and

$$\begin{pmatrix} 860 & -1025 \\ -1025 & 1228 \end{pmatrix}^{-1} = \begin{pmatrix} 0.2251146 & 0.1879010 \\ 0.1879010 & 0.1576535 \end{pmatrix}$$

$$\begin{pmatrix} 210.2 & -249.8 \\ -249.8 & 303.2 \end{pmatrix}^{-1} = \begin{pmatrix} 0.227525 & 0.187453 \\ 0.187453 & 0.157737 \end{pmatrix}$$

to fit a model of y on x_1 and x_2, i.e., do the following regression model,

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i, \quad \epsilon_i \overset{i.i.d.}{\sim} N(0, \sigma^2)$$

(a) **Assume that $\boldsymbol{\beta_0 = 2}$.**

i. Find the least squares estimates of the unknown parameters β_1 and β_2, then write down the fitted line.

ii. Find the Residual Sum of Squares and the unbiased estimate of the unknown parameter σ^2.

(b) **Assume that $2\boldsymbol{\beta_1 = \beta_2}$.**

· Find the least squares estimates of the unknown parameters β_1 and β_2, then write down the fitted line.

(c) **Assume that $\boldsymbol{\beta_0, \beta_1, \beta_2}$ are all unkown.**

· Find the least squares estimates of the unknown parameters β_0, β_1 and β_2, then write down the fitted line.

2. Consider a situation in which the regression data set is divided into two parts as follows. The model is given by

$$
\begin{aligned}
y_i &= \beta_0^{(1)} + \beta_1 x_i + \epsilon_i, \quad \text{for } i = 1, \cdots, n_1 \\
&= \beta_0^{(2)} + \beta_1 x_i + \epsilon_i, \quad \text{for } i = n_1 + 1, \cdots, n_1 + n_2
\end{aligned}
$$

In other words there are two regression lines with common slope. Using the centered model,

$$
\begin{aligned}
y_i &= \beta_0^{(1)^*} + \beta_1 (x_i - \bar{x}_1) + \epsilon_i, \quad \text{for } i = 1, \cdots, n_1 \\
&= \beta_0^{(2)^*} + \beta_1 (x_i - \bar{x}_2) + \epsilon_i, \quad \text{for } i = n_1 + 1, \cdots, n_1 + n_2
\end{aligned}
$$

where $\bar{x}_1 = \frac{1}{n_1} \sum_{i=1}^{n_1} x_i$ and $\bar{x}_2 = \frac{1}{n_2} \sum_{i=n_1+1}^{n_1+n_2} x_i$.

Show that the least squares estimate of β_1 is given by

$$
\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (x_i - \bar{x}_1) y_i + \sum_{i=n_1+1}^{n_1+n_2} (x_i - \bar{x}_2) y_i}{\sum_{i=1}^{n_1} (x_i - \bar{x}_1)^2 + \sum_{i=n_1+1}^{n_1+n_2} (x_i - \bar{x}_2)^2}
$$

Sol of Q2:

(a) (i) $y_i = 2 + \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i \quad i = 1, \cdots n$

$\Longleftrightarrow \underbrace{y_i - 2}_{y_i'} = \beta_1 x_{i1} + \beta_2 x_{i2} + \epsilon_i \quad i = 1, \cdots, n$

(no intercept model).

$$y' = \begin{pmatrix} y_1' \\ y_2' \\ \vdots \\ y_n' \end{pmatrix} = \begin{pmatrix} y_1 - 2 \\ y_2 - 2 \\ \vdots \\ y_n - 2 \end{pmatrix} \qquad X = [x_1, x_2] \in \mathbb{R}^{n \times 2}$$

$\uparrow$ column vector of $n \times 1$

$$X^T X = \begin{bmatrix} x_1^T \\ x_2^T \end{bmatrix} [x_1 \ x_2] = \begin{bmatrix} x_1^T x_1 & x_1^T x_2 \\ x_2^T x_1 & x_2^T x_2 \end{bmatrix} = \begin{bmatrix} \sum x_{i1}^2 & \sum x_{i1} x_{i2} \\ \sum x_{i1} x_{i2} & \sum x_{i2}^2 \end{bmatrix}$$

$$X^T y' = \begin{bmatrix} x_1^T \\ x_2^T \end{bmatrix} y' = \begin{bmatrix} x_1^T y' \\ x_2^T y' \end{bmatrix} = \begin{bmatrix} \sum x_{i1}(y_i - 2) \\ \sum x_{i2}(y_i - 2) \end{bmatrix} = \begin{bmatrix} \sum x_{i1} y_i - 2\sum x_{i1} \\ \sum x_{i2} y_i - 2\sum x_{i2} \end{bmatrix}$$

$$(X^T X)^{-1} = \begin{pmatrix} 860 & -1025 \\ -1025 & 1228 \end{pmatrix}^{-1} = \underline{\begin{pmatrix} 0.225 & 0.188 \\ 0.188 & 0.158 \end{pmatrix}} \qquad = \underline{\begin{bmatrix} 1309 \\ -1502 \end{bmatrix}}$$

$$\hat{\beta} = (X^T X)^{-1} X^T y' = \underline{\begin{pmatrix} 3.053 \\ 1.283 \end{pmatrix}} = \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{pmatrix}$$

$$\hat{y} = \underline{2} + 3.053 x_1 + 1.283 x_2$$

(ii) $\hat{\sigma}^2 = \dfrac{SSE}{n-p}$

$SSE = \sum_{i=1}^{n} (y_i' - \hat{y}_i')^2 \qquad e_i = y_i' - \hat{y}_i'$

$$e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} = (I - \underset{\uparrow \atop X(X^T X)^{-1} X}{H})\, y'$$

$$SSE = \|e\|^2 = e^T e = y'^T \underline{(I-H)}\,\underline{(I-H)}\, y'$$

$$= y'^T (I-H) y' = y'^T y' - \underline{y'}^T \underline{X} \underbrace{(X^T X)^{-1} X y'}_{\hat{\beta}}$$

$$= y'^T y' - (X\hat{\beta})^T y'$$

$$= y'^T y' - \hat{\beta}^T \underline{X^T y'}$$

$$= \underline{\sum_{i=1}^{n} (y_i - 2)^2} - 2004.778$$

$$= \sum_{i=1}^{n} y_i^2 - 4 \sum y_i + 4n - 2004.778$$

$$= 137.222$$

$$\hat{\sigma}^2 = \frac{SSE}{18} = 7.623$$

(b)

$$y_i = \beta_0 + \underline{\beta_1} (\underbrace{x_{i1} + 2x_{i2}}_{\underline{x_i'}}) + \epsilon_i \qquad i = 1, \cdots, n.$$

$$\bar{x'} = \bar{x}_1 + 2\bar{x}_2$$

$$\hat{\beta}_1 = \frac{S_{x'y}}{S_{x'x'}} = \frac{\sum (x_i' - \bar{x'})(y_i - \bar{y})}{\sum (x_i' - \bar{x'})^2}$$

$$= \frac{\sum (x_{i1} + 2x_{i2} - \bar{x}_1 - 2\bar{x}_2)(y_i - \bar{y})}{\sum \left[(x_{i1} - \bar{x}_1)^2 + 4(x_{i2} - \bar{x}_2)^2 + 4(x_{i1} - \bar{x}_1)(x_{i2} - \bar{x}_2) \right]}$$

$$= \frac{S_{x_1 y} + 2 S_{x_2 y}}{S_{x_1 x_1} + 4 S_{x_2 x_2} + 4 S_{x_1 x_2}} = -0.842850$$

$$\hat{\beta}_2 = 2\hat{\beta}_1$$

$$\hat{\beta}_0 = \bar{y} - (\hat{\beta}_1 \bar{x}_1 + \hat{\beta}_2 \bar{x}_2) = 4.44$$

(c) $\underline{y_i = \beta_0' + \beta_1 x_{i1}' + \beta_2 x_{i2}' + \epsilon_i}$ $\quad i = 1, \dots, n$

$$x_{i1}' = x_{i1} - \bar{x}_1, \quad x_{i2}' = x_{i2} - \bar{x}_2$$

$$X_c = [\underset{\downarrow}{1_n}, x_1', x_2']$$

Column vector of all 1's of length n

$$X_c^T X_c = \begin{bmatrix} 1_n^T \\ x_1'^T \\ x_2'^T \end{bmatrix} [1_n \; x_1' \; x_2'] = \begin{bmatrix} 1_n^T 1_n & 1_n^T x_1' & 1_n^T x_2' \\ 1_n^T x_1' & \sum x_{i1}'^2 & \sum x_{i1} x_{i2} \\ 1_n^T x_2' & \sum x_{i1}' x_{i2}' & \sum x_{i2}'^2 \end{bmatrix}$$

$$= \begin{bmatrix} n & 0 & 0 \\ 0 & S_{x_1 x_1} & S_{x_1 x_2} \\ 0 & S_{x_1 x_2} & S_{x_2 x_2} \end{bmatrix}$$

$$(X_c^T X_c)^{-1} = \begin{bmatrix} \frac{1}{n} & 0 & 0 \\ 0 & \multicolumn{2}{c}{\multirow{2}{*}{$\begin{pmatrix} \underline{S_{x_1x_1}} & \underline{S_{x_1x_2}} \\ S_{x_1x_2} & S_{x_2x_2} \end{pmatrix}^{-1}$}} \\ 0 & & \end{bmatrix}$$

$$X = f(X_c)$$

$$\underline{X_c (X_c^T X_c)^{-1} X_c^T}$$

$$\underline{X (X^T X)^{-1} X^T}$$

$$X_c^T y = \begin{bmatrix} \sum y_i \\ \sum x_{i1} y_i \\ \sum x_{i2} y_i \end{bmatrix}$$

$$X_c = f(X)$$

$$\hat\beta = (X_c^T X_c)^{-1} X_c^T y = \begin{pmatrix} \underline{11.1} \\ \underline{2.8606} \\ \underline{1.3197} \end{pmatrix} = \begin{pmatrix} \underline{\hat\beta_0'} \\ \underline{\hat\beta_1} \\ \underline{\hat\beta_2} \end{pmatrix}$$

$$\hat\beta_0 = \bar y - \hat\beta_1 \bar x_1 - \hat\beta_2 \bar x_2 = 3.76888$$

2.

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_{n_1} \\ y_{n_1+1} \\ \vdots \\ y_{n_1+n_2} \end{bmatrix} \quad X = \begin{bmatrix} 1 & 0 & x_1 - \bar{x}_1 \\ \vdots & \vdots & \vdots \\ 1 & 0 & x_{n_1} - \bar{x}_1 \\ 0 & 1 & x_{n_1+1} - \bar{x}_2 \\ \vdots & \vdots & \vdots \\ 0 & 1 & x_{n_1+n_2} - \bar{x}_2 \end{bmatrix} \quad \beta = \begin{pmatrix} \beta_0^{(1)*} \\ \beta_0^{(2)*} \\ \beta_1 \end{pmatrix}$$

$$\epsilon = \begin{bmatrix} \epsilon_1 \\ \vdots \\ \epsilon_{n_1+n_2} \end{bmatrix}$$

$$\underline{y = X\beta + \epsilon}$$

$$\hat{\beta} = (X^T X)^{-1} X^T y$$

$$X^T = \begin{bmatrix} 1 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 1 & \cdots & 1 \\ x_1 - \bar{x}_1 & \cdots & & & \cdots & x_{n_1+n_2} - \bar{x}_2 \end{bmatrix}_{3 \times (n_1 + n_2)}$$

$$X^T X = \begin{bmatrix} n_1 & 0 & 0 \\ 0 & n_2 & 0 \\ 0 & & \sum_{i=1}^{n_1} (x_i - \bar{x}_1)^2 + \sum_{i=n_1+1}^{n_1+n_2} (x_i - \bar{x}_2)^2 \end{bmatrix}$$

$$\underline{(X^T X)^{-1}} = \begin{bmatrix} \frac{1}{n_1} & & \\ & \frac{1}{n_2} & \\ & & \frac{1}{\sum_{i=1}^{n_1} (x_i - \bar{x}_1)^2 + \sum_{i=n_1+1}^{n_1+n_2} (x_i - \bar{x}_2)^2} \end{bmatrix}$$

$$X^T y = \begin{bmatrix} \sum_{i=1}^{n_1} y_i \\ \sum_{i=n_1+1}^{n_1+n_2} y_i \\ \sum_{i=1}^{n_1} (x_i - \bar{x}_1) y_i + \sum_{i=n_1+1}^{n_1+n_2} (x_i - \bar{x}_2) y_i \end{bmatrix}$$

$$\hat{\beta} = \begin{bmatrix} \frac{1}{n_1} \sum_{i=1}^{n_1} y_i \\ \frac{1}{n_2} \sum_{i=n_1+1}^{n_1+n} y_i \\ \hat{\beta}_1 \end{bmatrix}$$

$$(I - H)(I - H)$$

$$\underline{H^2 = H}$$

$$= I - H - \underline{H + H^2} = I - H$$