REGRESSION ANALYSIS - HOMEWORK 1 SOLUTIONS

ALAITTIN KIRTISOGLU

1) The least square function is $S(\beta_0, \beta_1) = \sum\limits_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)^2$. Set the partial derivatives of S to 0.

$$\frac{\partial S}{\partial \beta_0} = -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i) = 0, \qquad \frac{\partial S}{\partial \beta_1} = -2\sum_{i=1}^{n}(y_i - \beta_0 - \beta_1 x_i)x_i = 0. \tag{1}$$

Find the solutions $\hat{\beta}_0$, $\hat{\beta}_1$ of normal equations in (1) as follows.

$$\hat{\beta}_0 = \frac{\sum\limits_{i=1}^{n} y_i \sum\limits_{i=1}^{n} {x_i}^2 - \sum\limits_{i=1}^{n} x_i \sum\limits_{i=1}^{n} x_i y_i}{n\sum\limits_{i=1}^{n} {x_i}^2 - \left(\sum\limits_{i=1}^{n} x_i\right)^2}, \qquad \hat{\beta}_1 = \frac{n\sum\limits_{i=1}^{n} x_i y_i - \sum\limits_{i=1}^{n} x_i \sum\limits_{i=1}^{n} y_i}{n\sum\limits_{i=1}^{n} {x_i}^2 - \left(\sum\limits_{i=1}^{n} x_i\right)^2} \tag{2}$$

Note that arithmetic means are $\bar{x} = \frac{1}{n}\sum\limits_{i=1}^{n} x_i$ and $\bar{y} = \frac{1}{n}\sum\limits_{i=1}^{n} y_i$. Simplify $\hat{\beta}_1$ in (2) and find the following formula. Finally, we obtain $\hat{\beta}_0$ by substituting $\hat{\beta}_1$ in $\hat{\beta}_0$ in (2).

$$\hat{\beta}_1 = \frac{\sum\limits_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sum\limits_{i=1}^{n}(x_i - \bar{x})^2}, \qquad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}.$$

2.1) A regression line has two free parameters $\hat{\beta}_0$ and $\hat{\beta}_1$. But the line must pass through $(\bar{x}, \bar{y})$.

2.2) $\sum\limits_{i=1}^{n}(\hat{y}_i - \bar{y}) = \sum\limits_{i=1}^{n}\hat{\beta}_0 + \sum\limits_{i=1}^{n}\hat{\beta}_1 x_i - \sum\limits_{i=1}^{n}\bar{y} = \sum\limits_{i=1}^{n}\bar{y} - \sum\limits_{i=1}^{n}\hat{\beta}_1\bar{x} + \sum\limits_{i=1}^{n}\hat{\beta}_1 x_i - \sum\limits_{i=1}^{n}\bar{y} = \hat{\beta}_1 n\bar{x} + \hat{\beta}_1\sum\limits_{i=1}^{n} x_i = 0.$

2.3) $\sum\limits_{i=1}^{n}(y_i - \hat{y}_i) = \sum\limits_{i=1}^{n} y_i - \sum\limits_{i=1}^{n}\hat{y}_i = n\bar{y} - \sum\limits_{i=1}^{n}\hat{y}_i = \sum\limits_{i=1}^{n}\bar{y} - \sum\limits_{i=1}^{n}\hat{y}_i = -\sum\limits_{i=1}^{n}(\hat{y}_i - \bar{y}) = 0,$ **by 2.2.**

2.4) $\sum\limits_{i=1}^{n} x_i e_i = \sum\limits_{i=1}^{n} x_i(y_i - \hat{y}_i) = \sum\limits_{i=1}^{n} x_i y_i - \sum\limits_{i=1}^{n} x_i(\hat{\beta}_0 + \hat{\beta}_1 x_i) = \sum\limits_{i=1}^{n} x_i y_i - \sum\limits_{i=1}^{n}\hat{\beta}_1 x_i^2 - \sum\limits_{i=1}^{n} x_i\hat{\beta}_0$

$= \sum\limits_{i=1}^{n} x_i y_i - \sum\limits_{i=1}^{n}\hat{\beta}_1 x_i^2 - \sum\limits_{i=1}^{n} x_i(\bar{y} - \hat{\beta}_1\hat{x}) = \sum\limits_{i=1}^{n} x_i y_i - \sum\limits_{i=1}^{n}\hat{\beta}_1 x_i^2 - \sum\limits_{i=1}^{n} x_i\bar{y} + \sum\limits_{i=1}^{n}\hat{\beta}_1 x_i\bar{x} = 0.$

2.5) $\sum\limits_{i=1}^{n}\hat{y}_i e_i = \sum\limits_{i=1}^{n}(\hat{\beta}_0 + \hat{\beta}_1 x_i)e_i = \hat{\beta}_0\sum\limits_{i=1}^{n} e_i + \hat{\beta}_1\sum\limits_{i=1}^{n} x_i e_i$. By 2.3, $\hat{\beta}_0\sum\limits_{i=1}^{n} e_i = 0$. By 2.4, $\sum\limits_{i=1}^{n} x_i e_i = 0$.

2.6) $SST = \sum\limits_{i=1}^{n}(y_i - \bar{y})^2 = \sum\limits_{i=1}^{n}(y_i - \hat{y}_i + \hat{y}_i - \bar{y}_i)^2 = \sum\limits_{i=1}^{n}(y_i - \hat{y}_i)^2 + \sum\limits_{i=1}^{n}(\hat{y}_i - \bar{y})^2 + 2\sum\limits_{i=1}^{n}(y_i - \hat{y}_i)(\hat{y}_i - \bar{y})$

$= SSE + SSR + 2\sum\limits_{i=1}^{n}(y_i - \hat{y}_i)\hat{y}_i - 2\bar{y}\sum\limits_{i=1}^{n}(y_i - \hat{y}_i)$. By 2.5, $2\sum\limits_{i=1}^{n}(y_i - \hat{y}_i)\hat{y}_i = 0$. By 2.3, $\bar{y}\sum\limits_{i=1}^{n}(y_i - \hat{y}_i) = 0$.

Date: October 24, 2023.

4) Let $\mathbf{u} = (y_1 - \bar{y}, y_2 - \bar{y}, ...y_n - \bar{y})^T \in \mathbb{R}^n$ and $\mathbf{v} = ((x_1 - \bar{x}, x_2 - \bar{x}, ...x_n - \bar{x})^T \in \mathbb{R}^n$. Then,

. $$Cor(Y,X) = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2 \sum_{i=1}^{n}(x_i - \bar{x})^2}} = \frac{\mathbf{u}^T\mathbf{v}}{\|u\|\|v\|} = cos(\theta),$$ where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$.

`Questions 3 and 5 are in the following pages.`

Illinois Institute of Technology, Applied Mathematics, Chicago, IL, USA

Email address: akirtisoglu@hawk.iit.edu