Tutorial 01 answers

Patrick Li

2023-02-25

Exercise 1A: Some mathematical derivations

a. Show that $\sum_{i=1}^n (x_{i1} - \bar{x}_1)^2 = \sum_{i=1}^n x_{i1}^2 - n\bar{x}_1^2$

$$
\begin{aligned}
\sum_{i=1}^n (x_{i1} - \bar{x}_1)^2 &= \sum_{i=1}^n (x_{i1}^2 + \bar{x}_1^2 - 2x_{i1}\bar{x}_1) \\
&= \sum_{i=1}^n x_{i1}^2 + \sum_{i=1}^n \bar{x}_1^2 - \sum_{i=1}^n 2x_{i1}\bar{x}_1 \\
&= \sum_{i=1}^n x_{i1}^2 + n\bar{x}_1^2 - \sum_{i=1}^n 2x_{i1}\bar{x}_1 \\
&= \sum_{i=1}^n x_{i1}^2 + n\bar{x}_1^2 - 2\bar{x}_1 \sum_{i=1}^n x_{i1} \\
&= \sum_{i=1}^n x_{i1}^2 + n\bar{x}_1^2 - 2\bar{x}_1 n\bar{x}_1 \\
&= \sum_{i=1}^n x_{i1}^2 - n\bar{x}_1^2.
\end{aligned}
$$

Consider the regression $y_i = \beta_1 x_{i1} + e_i$, where the RSS is $\sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \beta_1 x_{i1})^2$.

b. Show that $\sum_{i=1}^n e_i^2 = \sum_{i=1}^n y_i^2 - 2\beta_1 \sum_{i=1}^n y_i x_{i1} + \beta_1^2 \sum_{i=1}^n x_{i1}^2$

$$
\begin{aligned}
\sum_{i=1}^n e_i^2 &= \sum_{i=1}^n (y_i - \beta_1 x_{i1})^2 \\
&= \sum_{i=1}^n (y_i^2 + \beta_1^2 x_{i1}^2 - 2y_i\beta_1 x_{i1}) \\
&= \sum_{i=1}^n y_i^2 + \sum_{i=1}^n \beta_1^2 x_{i1}^2 - \sum_{i=1}^n 2y_i\beta_1 x_{i1} \\
&= \sum_{i=1}^n y_i^2 + \beta_1^2 \sum_{i=1}^n x_{i1}^2 - 2\beta_1 \sum_{i=1}^n y_i x_{i1}.
\end{aligned}
$$

c. Using your answers to exercise 1A a. compute an expression for the derivative $\frac{\partial}{\partial \beta_1} \sum_{i=1}^n e_i^2$.

$$\begin{aligned}
\frac{\partial}{\partial \beta_1}\sum_{i=1}^{n} e_i^2 &= \frac{\partial}{\partial \beta_1}\left(\sum_{i=1}^{n} y_i^2 - 2\beta_1\sum_{i=1}^{n} y_i x_{i1} + \beta_1^2\sum_{i=1}^{n} x_{i1}^2\right)\\
&= -2\sum_{i=1}^{n} y_i x_{i1} + 2\beta_1\sum_{i=1}^{n} x_{i1}^2\\
&= 2\left(\beta_1\sum_{i=1}^{n} x_{i1}^2 - \sum_{i=1}^{n} y_i x_{i1}\right).
\end{aligned}$$

d. Find the parameter value $\hat{\beta}_1$ such that $\frac{\partial}{\partial \hat{\beta}_1}\sum_{i=1}^{n} e_i^2 = 0$.

Let

$$\begin{aligned}
\frac{\partial}{\partial \hat{\beta}_1}\sum_{i=1}^{n} e_i^2 &= 0\\
2\left(\hat{\beta}_1\sum_{i=1}^{n} x_{i1}^2 - \sum_{i=1}^{n} y_i x_{i1}\right) &= 0\\
\hat{\beta}_1\sum_{i=1}^{n} x_{i1}^2 &= \sum_{i=1}^{n} y_i x_{i1}\\
\hat{\beta}_1 &= \frac{\sum_{i=1}^{n} y_i x_{i1}}{\sum_{i=1}^{n} x_{i1}^2}.
\end{aligned}$$

Since $Cov(x_1, y) = (\sum_{i=1}^{n} x_{i1}y_i - n\bar{x}_1\bar{y})/(n-1)$ and $Var(x_1) = \sum_{i=1}^{n}(x_{i1} - \bar{x}_1)^2/(n-1) = (\sum_{i=1}^{n} x_{i1}^2 - n\bar{x}_1^2)/(n-1)$, $\hat{\beta}_1$ can also be written as

$$\begin{aligned}
\hat{\beta}_1 &= \frac{\sum_{i=1}^{n} y_i x_{i1}}{\sum_{i=1}^{n} x_{i1}^2 - n\bar{x}_1^2 + n\bar{x}_1^2}\\
&= \frac{\sum_{i=1}^{n} y_i x_{i1}}{\sum_{i=1}^{n}(x_{i1}^2 - \bar{x}_1)^2 + n\bar{x}_1^2}\\
&= \frac{\sum_{i=1}^{n} y_i x_{i1}}{(n-1)Var(x_1) + n\bar{x}_1^2}\\
&= \frac{\sum_{i=1}^{n} y_i x_{i1} - n\bar{x}_1\bar{y} + n\bar{x}_1\bar{y}}{(n-1)Var(x_1) + n\bar{x}_1^2}\\
&= \frac{(n-1)Cov(x_1, y) + n\bar{x}_1\bar{y}}{(n-1)Var(x_1) + n\bar{x}_1^2}\\
&= \frac{Cov(x_1, y) + \frac{n}{n-1}\bar{x}_1\bar{y}}{Var(x_1) + \frac{n}{n-1}\bar{x}_1^2}.
\end{aligned}$$