Data100 hw5

Shangheng Chen

January 2021

1 Properties of Simple Linear Regression

(a)

$$e_i = y_i - \hat{y_i} = y_i - \bar{y} - r\sigma_y \frac{x_i - \bar{x}}{\sigma_x}$$

$$\sum_{i=1}^{n} e_i = \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} \bar{y} - \sum_{i=1}^{n} r\sigma_y \frac{x_i - \bar{x}}{\sigma_x}$$

$$\sum_{i=1}^{n} e_i = n\bar{y} - n\bar{y} - r\frac{\sigma_y}{\sigma_x} \sum_{i=1}^{n} x_i - \bar{x} = 0$$

(b)

since

$$\sum_{i=1}^{n} e_i = 0$$

then

$$\sum_{i=1}^{n} y_i - \hat{y_i} = 0$$

$$\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} \hat{y_i}$$

$$n\bar{y_i} = n\bar{\hat{y_i}}$$

$$\bar{y_i} = \bar{\hat{y_i}}$$

(c)

since

$$\hat{y} = \bar{y} + r\sigma_y \frac{x - \bar{x}}{\sigma_x}$$

use $x = \bar{x}$ to substitute x, we get

$$\hat{y} = \bar{y}$$

So, $(\bar{x}, \bar{y})$ is on the simple linear regression line.

2 Geometric Perspective of Least Squares

(a) From Lecture, we know that $\mathbb{Y} - \hat{\mathbb{Y}}$ is orthogonal to the plane. So, it is also orthogonal to $\vec{1}$. Then,

$$\vec{1} \cdot e = 0$$

So,

$$\sum_{i=1}^{n} e_i = 0$$

(b) In order to minimize the norm of e, e should be orthogonal to the plane. So, $\vec{x}$ and vector e are orthogonal.
(c) We know that $\hat{\mathbb{Y}} = \hat{\theta_0}\vec{1} + \hat{\theta_1}\vec{x}$. Both $\vec{1}$ and $\vec{x}$ are orthogonal to e. Then $\hat{\mathbb{Y}}$ is orthogonal to e.

3 Properties of a Linear Model With No Constant Term

3.

$$\frac{dR(\gamma)}{d\gamma} = -\frac{2}{n}\sum_{i=1}^{n}(y_i - \gamma x_i)x_i$$

Let the derivative to zero, we get

$$\sum_{i=1}^{n} x_i y_i = \gamma \sum_{i=1}^{n} {x_i}^2$$

Then

$$\hat{\gamma} = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} {x_i}^2}$$

4.
(a)False
(b)True
(c)False
(d)False