Stat 451
11-17-16

① 

Find the best line

Model: $i=1,\dots,n \qquad y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$

②

Assume that $\varepsilon_1, \dots, \varepsilon_n$ are independent, identically distributed $N(0, \sigma^2)$

Then $y_1, \dots, y_n$ are independent, with $y_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2)$

$$\text{lik}(\beta_0, \beta_1, \sigma) = \prod_{i=1}^{n} f(y_i)$$

$$= \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y_i - \beta_0 - \beta_1 x_i}{\sigma}\right)^2}$$

3

$$\frac{\partial \ell}{\partial \beta_1} = -\frac{\cancel{1}}{\cancel{2}\sigma^2} \sum \cancel{2}(y_i - \beta_0 - \beta_1 x_i)(\cancel{-}x_i) \overset{\text{set}}{=} 0$$

$$\sum y_i x_i - \beta_0 \sum x_i - \beta_1 \sum x_i^2 = 0$$

substitute

$$\sum x_i y_i - (\bar{y} - \beta_1 \bar{x}) \sum x_i - \beta_1 \sum x_i^2 = 0$$

$$\sum x_i y_i - \bar{y} \sum x_i + \beta_1 \bar{x} \sum x_i - \beta_1 \sum x_i^2 = 0$$

$$\beta_1 \left(\sum x_i^2 - \bar{x} \sum x_i \right) = \sum x_i y_i - \bar{y} \sum x_i$$

$$= \sigma^{-n} (2\pi)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2}$$ 4

$$\ell(\beta_0, \beta_1, \sigma) = -n \ln \sigma - \frac{n}{2} \ln(2\pi) - \frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

$$\frac{\partial \ell}{\partial \beta_0} = -\frac{\cancel{1}}{\cancel{2}\sigma^2} \sum \cancel{2}(y_i - \beta_0 - \beta_1 x_i)(\cancel{-1}) \overset{\text{set}}{=} 0$$

$$\sum y_i - n\beta_0 - \beta_1 \sum x_i = 0$$

$$\hat{\beta}_0 = \frac{\sum y_i}{n} - \beta_1 \frac{\sum x_i}{n}$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

5

$$\hat{\beta}_1 = \frac{\sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n}}{\sum x_i^2 - \frac{(\sum x_i)^2}{n}} \begin{matrix} \}\, SS_{xy} \\ \}\, SS_{xx} \end{matrix}$$

$$\hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}}, \quad \hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\frac{\partial \ell}{\partial \sigma} = -\frac{1}{\sigma} - \frac{1}{2}(-2\sigma^{-3}) \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 \overset{set}{=} 0$$

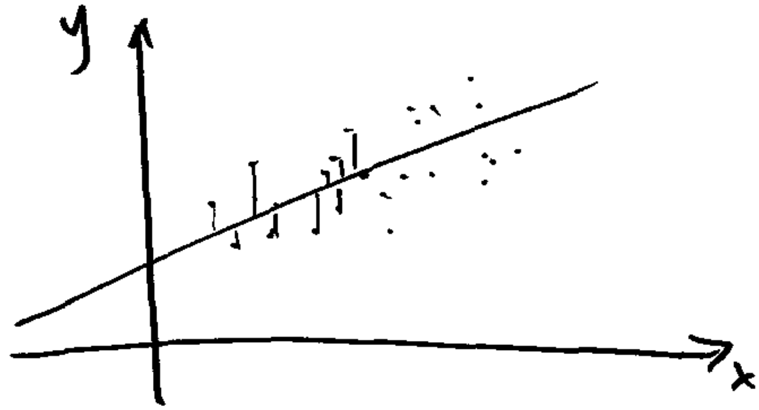
$$-n\sigma^2 + \sum (y_i - \beta_0 - \beta_1 x_i)^2 = 0$$

6

$$\hat{\sigma}^2 = \frac{\sum (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2}{n}$$

(7)

Least-squares solution:



Find the line that minimized the sum of squares of the <u>vertical distances</u> from the points to the line ↳ <u>residuals</u>

(8)

Model : $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$

No further assumptions

$$\varepsilon_i = y_i - \beta_0 - \beta_1 x_i$$

$$\sum_{i=1}^{n} \varepsilon_i^2 = \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2$$

$\parallel$

SSE

We want to minimize SSE

$$\frac{\partial SSE}{\partial \beta_0} = \sum 2(y_i - \beta_0 - \beta_1 x_i)(-1) \stackrel{set}{=} 0$$

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\frac{\partial SSE}{\partial \beta_1} = \sum 2(y_i - \beta_0 - \beta_1 x_i)(-x_i) \stackrel{set}{=} 0$$

$$\hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}}$$

Same solutions as we get using MLE !!