Ch8 簡單直線迴歸

1. 成對的資料 $(x_1, y_1)(x_2, y_2) \cdots (x_n, y_n)$ 欲找一條直線，模擬 X與Y之關係。

與 r 高度相關

2. 設 $Y = ax + b$

$$\boxed{\begin{aligned} \hat{a} &= \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \\ \hat{b} &= \bar{y} - a\bar{x} \end{aligned}}$$

截距

以 X 來預估 Y

$\hat{Y} = \hat{a}x + \hat{b}$ 稱為 Y 對 X 之回歸線。regression

Y 對 X 的直線。

3、a與b之值是由 $\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2$ 最小時得到的。

(例)

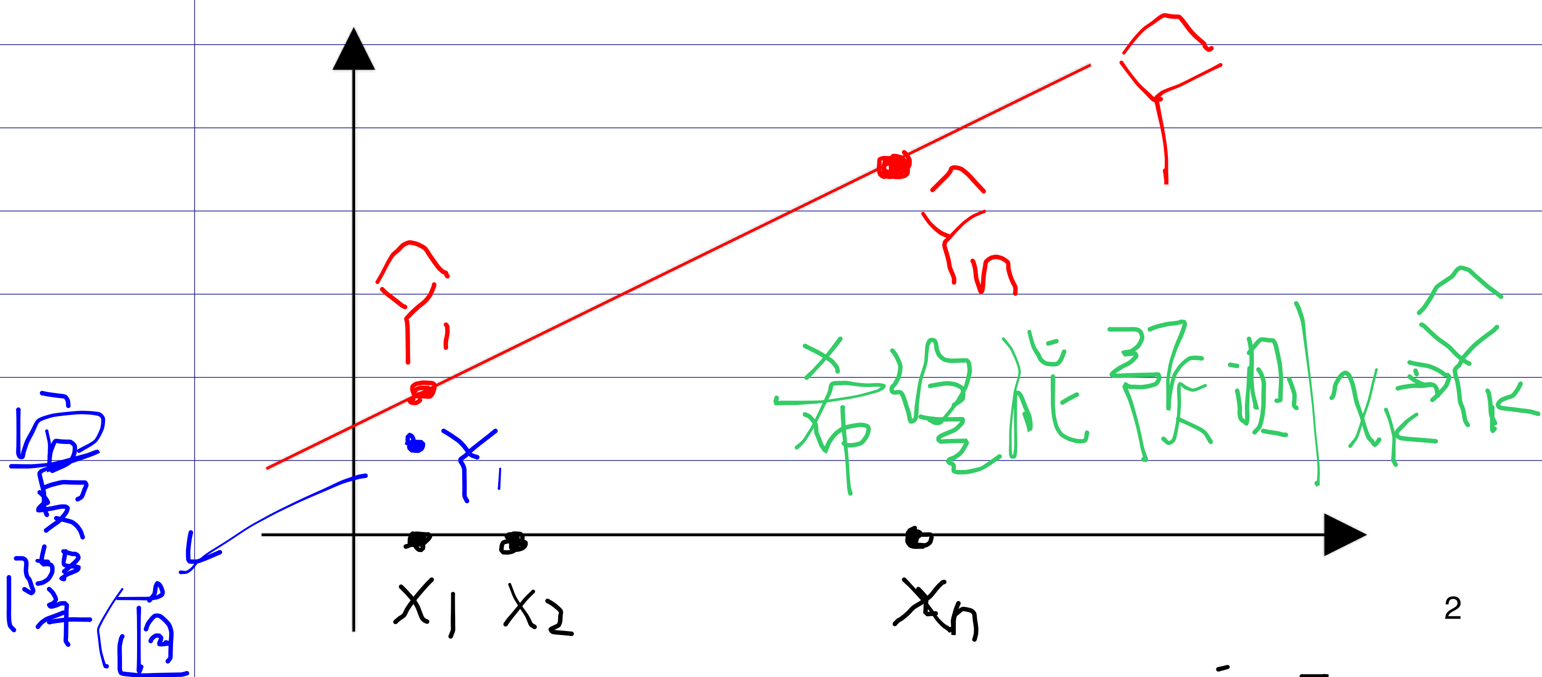
$$M = \quad = \sum_{i=1}^{n}(Y_i - aX_i - b)^2$$

展開後,將 a′, b′

偏微分

$$\frac{\partial M}{\partial a} \overset{令}{=} 0$$

$$\frac{\partial M}{\partial b} \overset{令}{=} 0$$



4. $\hat{a} > 0 \iff r > 0$

5. 資料標準化之後. $\hat{z}_2 = r z_1$

pf: 設 $\hat{z}_2 = \hat{a} z_1 + \hat{b}$

$$\hat{a} = \frac{\sum_{i=1}^{n} (z_{1i} \cdot z_{2i})}{\sum_{i=1}^{n} z_{1i}}$$

$$= \frac{\sum_{i=1}^{n} z_{1i} \cdot z_{2i}}{\sqrt{\sum_{i=1}^{n} z_{1i}^2 \cdot \sum_{i=1}^{n} z_{2i}^2}} \times \frac{\sqrt{\frac{\sum_{i=1}^{n} z_{2i}^2}{n-1}}}{\sqrt{\frac{\sum_{i=1}^{n} z_{1i}^2}{n-1}}} = r$$

$$\hat{b} = \bar{z}_2 - \hat{a} \bar{z}_1 = 0$$

$$\therefore \hat{z}_2 = r \hat{z}_1$$