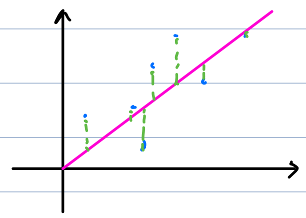
Regression Line

(x_i, y_i) , $i = 1, \cdots, n$

line $\quad y = \hat{a}x + \hat{b}$: best mean-square fit



minimizing $\sum_{i=1}^{n} (y_i - ax_i - b)^2$

$$F(a,b) = \sum_{i=1}^{n} (ax_i + b - y_i)^2$$

$$\frac{\partial F}{\partial a} = 2 \sum_{i=1}^{n} x_i (ax_i + b - y_i) = 0 \Rightarrow a \underbrace{\sum x_i^2}_{W_x} + bn\bar{x} = \underbrace{\sum x_i y_i}_{W_{xy}}$$

$$\frac{\partial F}{\partial b} = 2 \sum_{i=1}^{n} (ax_i + b - y_i) = 0 \Rightarrow an\bar{x} + nb = n\bar{y}$$

$$\begin{cases} W_x a + n\bar{x} b = W_{xy} \\ n\bar{x} a + nb = n\bar{y} \end{cases}$$

$$\Rightarrow \begin{cases} W_x a + n\bar{x} b = W_{xy} \\ -n\bar{x}^2 a - n\bar{x} b = -n\bar{x}\bar{y} \end{cases}$$

$$\Rightarrow \quad \hat{a} = \frac{W_{xy} - n\bar{x}\bar{y}}{W_x - n\bar{x}^2} \qquad \hat{b} = \bar{y} - \bar{x}\hat{a}$$

$$\frac{\partial^2 F}{\partial a^2} = \underbrace{2W_x}_{>0} \qquad \frac{\partial^2 F}{\partial a \partial b} = 2n\bar{x}$$

$$\det = 4n\underbrace{(W_x - n\bar{x}^2)}_{= n\sigma_x^2} > 0$$

$$\frac{\partial^2 F}{\partial a \partial b} = 2n\bar{x} \qquad \frac{\partial^2 F}{\partial b^2} = 2n$$

Regression line

1) $Y_i = aX_i + b + \varepsilon_i$ ε_i - measurement errors

$\Downarrow$

SLR simple linear regression

2) (X_i, Y_i) – i.i.d. from joint Gaussian

$\Downarrow$

NCT

$\bar{X} = \frac{1}{n} \sum X_i$

$\tilde{X}_i = X_i - \bar{X}$

$\bar{Y}$, $\tilde{Y}_i$ similar

$W_x = \sum X_i^2$ $\tilde{W}_x = \sum \tilde{X}_i^2$

$\hat{\alpha} = \dfrac{\sum \tilde{X}_i \tilde{Y}_i}{\tilde{W}_x}$ $\hat{b} = \bar{Y} - \hat{\alpha}\bar{X}$

$Y_i = aX_i + b + \varepsilon_i$

$$\begin{cases} \hat{a} - a = \dfrac{\sum \tilde{X}_i \varepsilon_i}{\tilde{W}_x} \\ \hat{b} - b = \bar{\varepsilon} - \dfrac{\frac{1}{n}\sum (X_i - \bar{X})\varepsilon_i}{\tilde{W}_x} \end{cases}$$

SLR theorem

1) if $\mathbb{E}|\varepsilon_i| < \infty$ and $\mathbb{E}\varepsilon_i$ is the same for all i,
then $\mathbb{E}\hat{a} = a$

2) if also $\mathbb{E}\varepsilon_i = 0$,
then $\mathbb{E}\hat{b} = b$

3) if also $\mathbb{E}\varepsilon_i = \sigma^2$ and $\mathbb{E}\varepsilon_i\varepsilon_j = 0$ $i \neq j$

then $MSE(\hat{a}) = \dfrac{\sigma^2}{\hat{w}_x}$ $MSE(\hat{b}) = \dfrac{\sigma^2 W_x}{n\,\tilde{w}_x}$

$cov(\hat{a}, \hat{b}) = -\dfrac{\sigma^2 \bar{x}}{\tilde{w}_x}$

4) if also ε_i are i.i.d. $N(0, \sigma^2)$,
then the rest

(three-parameter example for chapter 8-9) a, b, σ^2

$n-2$ df