

선형회귀는 두데이터간에 정확한수식 한변수가 변할 때 다른값을 어떤값이 나올지
 $y=ax+b$ 관계 $y=ax+b+e$ 일차식 원래는 $y=ax+b$ 라는 아름다운 식이 있었다.
 회귀는 원래대로 돌아간다는 뜻이다 본질로 그래서 에러를 제거한후 두데이터간의 관계를 도
 출한다.

선형회귀 유도 (기울기)

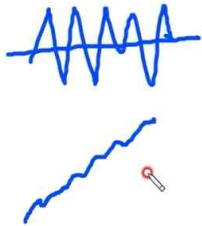
$$\begin{aligned}
 y &= ax + b + \epsilon \\
 \text{cov}(x, y) &= \text{cov}(x, ax + b + \epsilon) \\
 &= \text{cov}(x, ax) + \text{cov}(x, b) + \text{cov}(x, \epsilon) \\
 &= a \text{cov}(x, x) = aV(x) \\
 a &= \frac{\text{cov}(x, y)}{V(x)}
 \end{aligned}$$

Handwritten notes in red:
 - Above ϵ in the first equation: ϵ_{error}
 - Next to $\text{cov}(x, b)$: 0
 - Next to $\text{cov}(x, \epsilon)$: 0 변수 error
 - Under $a \text{cov}(x, x)$: a key icon

위에서 보다시피 $\text{cov}(x, e)$ 는상관계수라서 변수가 에러가 될 수 없고
 $\text{cov}(x, b)$ 는 0이라서 둘다 날아간다.

$\text{cov}(x, x)$ 자기자신이라서 분산이다.

선형회귀 유도 (y 절편)



$$\begin{aligned} \textcircled{y} &= ax + b + \epsilon \\ E(y) &= E(\cancel{ax} + b + \epsilon) \\ &= E(ax) + E(b) + E(\epsilon) \\ &= aE(x) + b \end{aligned}$$

$$\frac{\overbrace{b + b + \dots + b}^{N \text{ times}}}{N} = b$$

$$b = E(y) - aE(x)$$

$$= E(y) - \frac{\text{cov}(x, y)}{V(x)} E(x)$$

$$y = ax + b$$

$$y = \frac{\overset{\vee}{\text{cov}(x, y)}}{V(x)} x + E(y) - \frac{\text{cov}(x, y)}{V(x)} E(x)$$

$X \quad Y$

문제) $[1, 2, 3]$ 와 $[1, 3, 5]$ 의 선형회귀식?

$$E(Y) = 2 \quad \text{Cov}(X, Y) = \frac{4}{3}$$

$$E(X) = 3 \quad V(X) = \frac{2}{3}$$

abantu.code@gmail.com

$$Y = \frac{\frac{4}{3}}{\left(\frac{2}{3}\right)} X + 2 - \frac{\left(\frac{4}{3}\right)}{\left(\frac{2}{3}\right)} \cdot 2$$

$$= 2X + 3 - 4 = 2X - 1$$