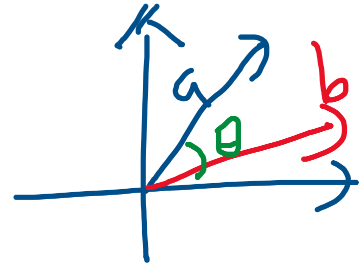
데이터구조와컴퓨팅 12주차

❑ 한준희

내적

$a=\begin{pmatrix}a_1\\a_2\end{pmatrix}$ $\quad b=\begin{pmatrix}b_1\\b_2\end{pmatrix}$



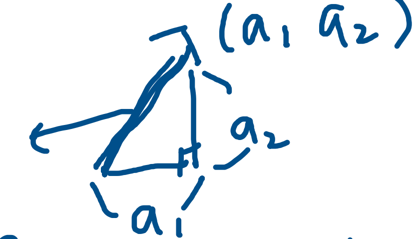
dot product

$a^t b = (a_1\ \ a_2)\begin{pmatrix}b_1\\b_2\end{pmatrix} = a_1b_1+a_2b_2 = a\cdot b$

inner product

$= |a||b|\cos\theta$

$\sqrt{a_1^2+a_2^2}$



$a\cdot b = 0 \iff \cos\theta = 0 \iff \theta = 90^\circ \iff a \perp b$

① $a\cdot a = a_1^2+a_2^2 = |a|^2$

$= |a||a|\underbrace{\cos 0^\circ}_{=1} = |a|^2$

$|a| = \sqrt{a\cdot a}$

② $a\cdot b = b\cdot a$

$$a\cdot b = |a||b|\cos\theta$$

$$|a| = \sqrt{a\cdot a}$$

$$|b| = \sqrt{b\cdot b}$$

$$\Rightarrow \cos\theta = \frac{a\cdot b}{|a||b|}$$

$$= \frac{a\cdot b}{\sqrt{a\cdot a}\sqrt{b\cdot b}}$$

6.123234e-17

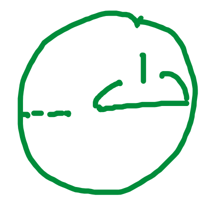
$$= 6.123234\times 10^{-17}$$

$$= \frac{6.\cdots}{\underbrace{1000\cdots00}_{17개}} \approx 0$$

각도 { degree, radian (라디안) 호도법

360°

2π



$$2\pi = 360°$$

$$\pi = 180°$$

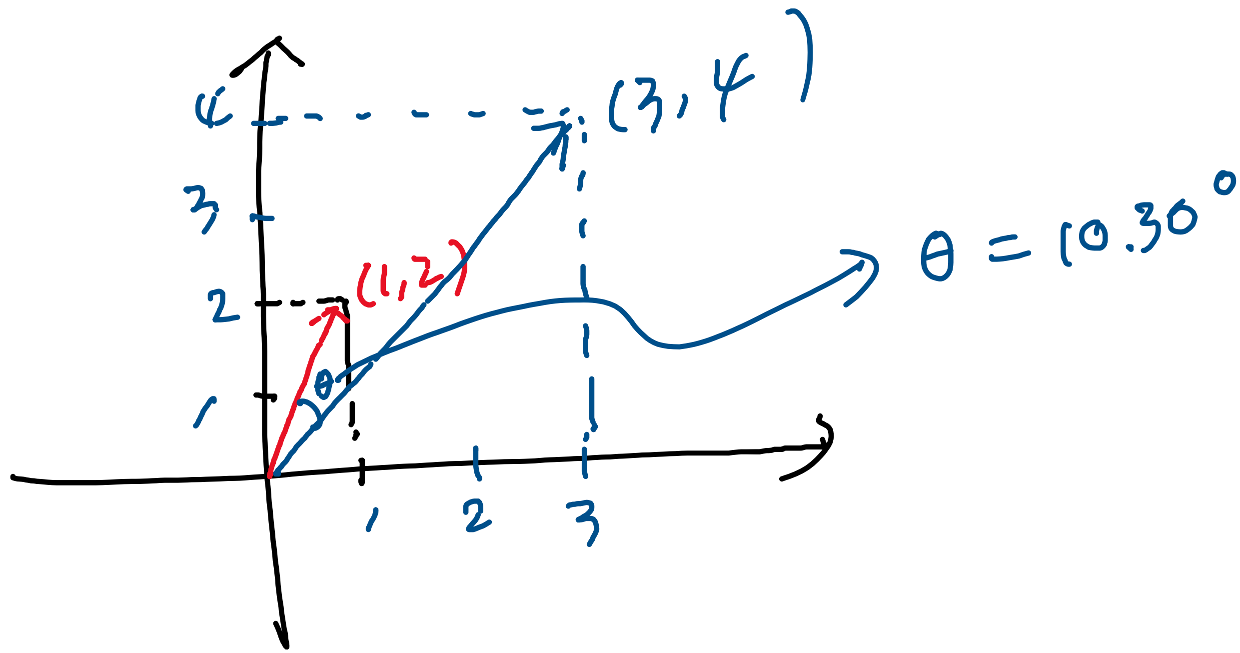
$$\frac{\pi}{2} = 90°$$

$$1° = \frac{\pi}{180°}$$

$$45° = \frac{\pi}{4}$$

$$\frac{180°}{\pi}$$

$$?° = 1$$

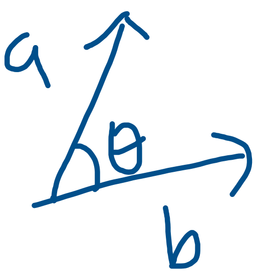

(3, 4)
(1, 2)
θ
θ = 10.30°
1
2
3
4
1
2
3

상관계수와 내적

$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ $y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ $-1 \le r \le 1$

데이터가 $(x_1, y_1), \dots, (x_n, y_n)$ 와 같이 주어졌을 때, 상관계수 r,

$a \cdot b = |a||b|\cos\theta$



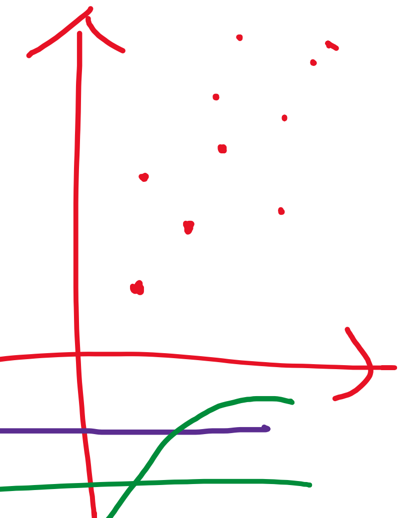
$$r = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2}\sqrt{\sum(y_i - \bar{y})^2}}$$

$a = x - \bar{x} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} x_1 - \bar{x} \\ \vdots \\ x_n - \bar{x} \end{pmatrix}$

$b = y - \bar{y} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} y_1 - \bar{y} \\ \vdots \\ y_n - \bar{y} \end{pmatrix}$

$$= \frac{\sum a_i b_i}{\sqrt{\sum a_i^2}\sqrt{\sum b_i^2}}$$

$\sum a$



$$= \frac{a \cdot b}{|a||b|} = \cos\theta \le 1$$

$-1 \le$

회귀분석과 행렬

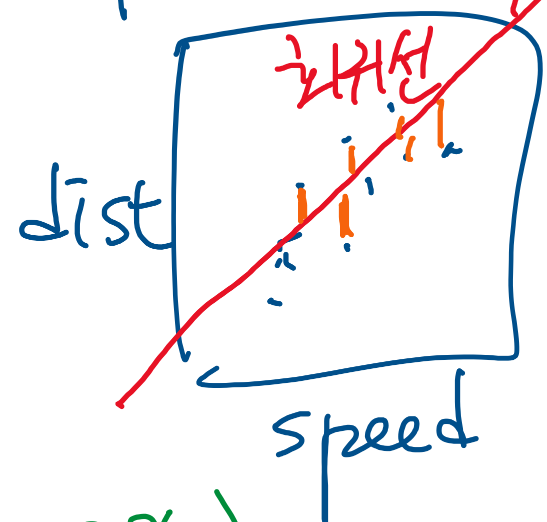
$\hat{\beta}_0$ $\hat{\beta}_1$ 회귀식

$$\text{dist} = -17.579 + 3.932 \times \text{speed}$$

$$y = \beta_0 + \beta_1 x + \epsilon$$

$$y = X\beta + \epsilon$$

$$\beta = \begin{pmatrix}\beta_0 \\ \beta_1\end{pmatrix}$$



$$\begin{pmatrix}1 & x_1 \\ 1 & x_2 \\ 1 & x_3\end{pmatrix}$$

$$y = \begin{pmatrix}y_1 \\ y_2 \\ y_3 \\ \vdots\end{pmatrix}$$

$$\begin{pmatrix}1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \\ \vdots & \end{pmatrix}\begin{pmatrix}\beta_0 \\ \beta_1\end{pmatrix} = \begin{pmatrix}\beta_0 + \beta_1 x_1 \\ \beta_0 + \beta_1 x_2 \\ \vdots\end{pmatrix}$$

$$y = X\hat{\beta}$$

$$X\hat{\beta} = y \quad \cdots\to \quad \hat{\beta} = X^{-1}y$$

문제!!

!!??

회귀분석과 행렬

$y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ ↳ 종속변수

$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ → 설명변수

$\beta = \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix}$

β_0 → y-절편

β_1 → 기울기

$y_1 = \beta_0 + \beta_1 x_1 + \varepsilon_1$

$y_2 = \beta_0 + \beta_1 x_2 + \varepsilon_2$

ε → 엡실론 (epsilon)

$\sum \varepsilon_i^2 = |\varepsilon|^2$

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}$$

$$y = X\beta + \varepsilon$$

회귀분석 과정

(ε_i들의 제곱의 합을 '최소화' 시키는 β를 구하는 것

최소화 → 미분

$$y = X\hat{\beta}$$

→ beta hat (추정값 = estimate)

한림대학교 HALLYM UNIVERSITY

회귀분석과 행렬

$$X^{-1}y = X^{-1}X\hat{\beta}$$

$$\hat{\beta} = X^{-1}y \quad ??\ \text{No!!}$$

$$(X^tX)^{-1}X^ty = \underbrace{(X^tX)^{-1}(X^tX)}_{I}\hat{\beta}$$

$$\therefore\ \hat{\beta} = (X^tX)^{-1}X^ty$$

$$X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ 1 & x_3 \end{pmatrix} \ \boxed{3\times2}$$

→ design (model) matrix

역행렬은 '정방행렬' 에대해서만 정의!

$\underset{2\times3\ \ 3\times2}{X^tX} \rightarrow \underline{2\times2}$

$\underset{3\times2\ \ 2\times3}{XX^t} \rightarrow 3\times3$

한림대학교
HALLYM UNIVERSITY