Doğukan Yıldırım 28364 [redacted]

(2)

(a)

$\hat{y}_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} = f(x_i)$ where x_i is the i^{th} sample $[x_{1i}, x_{2i}]$

3 samples

$$RSS = \sum_{i=1}^{N} (y_i - f(x_i))^2 = \sum_{i=1}^{N} (y_i - (\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}))^2$$

(b)

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (y_i - f(x_i))^2$$

$$\frac{\partial MSE}{\partial \beta_j} = \frac{\partial}{\partial \beta_j} \left(\frac{1}{N} \sum_{i=1}^{N} (y_i - (\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}))^2 \right)$$

$$= \frac{1}{N} \sum_{i=1}^{N} \frac{\partial}{\partial \beta_j} (y_i - (\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}))^2$$

$$= \frac{1}{N} \sum_{i=1}^{N} 2 (y_i - (\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}))(-x_{ji})$$

(c)

$$\beta_0 = \beta_0 - 0.5 \cdot \frac{2}{3} \sum_{i=1}^{3} (y_i - (\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i})(-x_{0i})$$

$$= -1 - \frac{1}{3} \cdot ((-1)\cdot(-1) + (-1)\cdot(-1) + (-1)\cdot(-1)) = -2$$

$$\beta_1 = \beta_1 - 0.5 \cdot \frac{2}{3} \sum_{i=1}^{3} (y_i - (\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}))(-x_{1i})$$

$$= 0 - \frac{1}{3} \cdot ((-1)\cdot(-1) + (-1)\cdot 0 + (-1)\cdot(0)) = -\frac{1}{3}$$

$$y_1 - (\beta_0 + \beta_1 x_{11} + \beta_2 x_{21}) = -2 - (-1 + 0 + 1 \cdot 0) = -1$$

$$y_2 - (\beta_0 + \beta_1 x_{12} + \beta_2 x_{22}) = 1 - (-1 + 3) = -1$$

$$y_3 - (\beta_0 + \beta_1 x_{13} + \beta_2 x_{23}) = 0 - (-1 + 2) = -1$$

Doğukan Yıldırım 28364

$$\beta_2 = \beta_2 - 0.5 \frac{2}{3} \sum_{i=1}^{3} (y_i - (\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i}))(-x_{2i})$$

$$= 1 - \frac{1}{3}((-1)\cdot 0 + (-1)(-3) + (-1)(-2)) = 1 - \frac{5}{3} = -\frac{2}{3}$$

Resulting weight vector: $[-2, -1/3, -2/3]^T$

New RSS.

$$y_1 - (\beta_0 + \beta_1 x_{11} + \beta_2 x_{21}) = -2 - (-2 - \tfrac{1}{3}\cdot 1 - \tfrac{2}{3}\cdot 0) = \tfrac{1}{3}$$

$$y_2 - (\beta_0 + \beta_1 x_{12} + \beta_2 x_{22}) = 1 - (-2 - \tfrac{1}{3}\cdot 0 - \tfrac{2}{3}\cdot 3) = 5$$

$$y_3 - (\beta_0 + \beta_1 x_{13} + \beta_2 x_{23}) = 0 - (-2 - \tfrac{1}{3}\cdot 0 - \tfrac{2}{3}\cdot 2) = \tfrac{10}{3}$$

$$RSS = (\tfrac{1}{3})^2 + (5)^2 + (\tfrac{10}{3})^2 \sim 36.23$$

Doğukan Yıldırım 28364 ~~[illegible]~~

(3)

$$\hat{\theta}_{MLE} = \underset{\theta}{\operatorname{argmax}}\ L(\theta) = \underset{\theta}{\operatorname{argmax}} \left(\log(L(\theta)) \right)$$

$$L(\theta) = \left(\frac{2\theta}{3}\right)^3 \cdot \left(\frac{\theta}{3}\right) \cdot (1-\theta)^2 \quad \text{since} \quad (y_1, y_2, y_3, y_4, y_5, y_6) = (1,2,3,3,1,1)$$

$$= \prod_{i=1}^{n} P(Y_i \mid \theta)$$

$$\log(L(\theta)) = \log\left(\prod_{i=1}^{n} P(Y_i \mid \theta)\right) = \log\left(\left(\frac{2\theta}{3}\right)^3 \cdot \left(\frac{\theta}{3}\right) \cdot (1-\theta)^2\right)$$

$$= \log\left(\frac{2\theta}{3}\right)^3 + \log\left(\frac{\theta}{3}\right) + \log(1-\theta)^2$$

$$= 3\log\left(\frac{2\theta}{3}\right) + \log\left(\frac{\theta}{3}\right) + 2\log(1-\theta)$$

$$= 3\log\left(\frac{2}{3}\right) + 3\log\theta + \log\left(\frac{1}{3}\right) + \log\theta + 2\log(1-\theta)$$

$$= 4\log\theta + 2\log(1-\theta) + 3\log\left(\frac{2}{3}\right) + \log\left(\frac{1}{3}\right)$$

$$\frac{d}{d\theta}\log(L(\theta)) = \frac{4}{\theta} - \frac{2}{1-\theta} = 0$$

$$\frac{4}{\theta} = \frac{2}{1-\theta} \Rightarrow 4(1-\theta) = 2\theta \Rightarrow 4 = 6\theta$$

$$\Rightarrow \hat{\theta}_{MLE} = \frac{2}{3}$$

Doğukan Yıldırım 28364

④

(a) P(at least one boy gets a book) = 1 − P(none of the boys get a book)

$$= 1 - \frac{20!\ 21!}{41!}$$

(b) A: randomly drawn ball is white

B: randomly drawn ball is numbered 2012

$$\left.\begin{matrix} 2 \text{ white} \\ 7 \text{ black} \end{matrix}\right) 2012$$

$$\left.\begin{matrix} 3 \text{ white} \\ 5 \text{ black} \end{matrix}\right) 2016$$

$$\left.\begin{matrix} 5 \text{ white} \\ 2 \text{ black} \end{matrix}\right) 2020$$

+

24 balls

A and B are independent $\Longleftrightarrow P(A|B) = P(A)$

$$P(A) = \frac{2+3+5}{24} = \frac{10}{24}$$

$$P(B) = \frac{2+7}{24} = \frac{9}{24}$$

$$P(A \cap B) = \frac{2}{24}$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{2}{24}}{\frac{9}{24}} = \frac{2}{9}$$

$P(A|B) = \frac{2}{9} \neq \frac{10}{24} = P(A) \Rightarrow$ A and B are NOT independent

Doğukan Yıldırım 28364

(5)

(a)

	X_1	X_2	X_3	X_4	X_5	Y
Observations	Volatile	Dividend	MACD_CONV	MACD_DIV	MACD_LNR	Risk
1	1	0	1	0	0	High
2	1	1	0	1	0	High
3	0	0	0	1	0	High
4	0	1	1	0	0	Low
5	0	1	0	0	1	Low
6	1	0	0	0	1	High

one-hot-encoded table

$P(X_1=1 \mid High) = \frac{3}{4}$ $\quad$ $P(X_1=1 \mid Low) = 0$

$P(X_1=0 \mid High) = \frac{1}{4}$ $\quad$ $P(X_1=0 \mid Low) = 1$

$P(X_2=1 \mid High) = \frac{1}{4}$ $\quad$ $P(X_2=1 \mid Low) = 1$

$P(X_2=0 \mid High) = 3/4$ $\quad$ $P(X_2=0 \mid Low) = 0$

$P(X_3=1 \mid High) = 1/4$ $\quad$ $P(X_3=1 \mid Low) = 1/2$

$P(X_3=0 \mid High) = 3/4$ $\quad$ $P(X_3=0 \mid Low) = 1/2$

$P(X_4=1 \mid High) = 1/2$ $\quad$ $P(X_4=1 \mid Low) = 0$

$P(X_4=0 \mid High) = 1/2$ $\quad$ $P(X_4=0 \mid Low) = 1$

$P(X_5=1 \mid High) = 1/4$ $\quad$ $P(X_5=1 \mid Low) = 1/2$

$P(X_5=0 \mid High) = 3/4$ $\quad$ $P(X_5=0 \mid Low) = 1/2$

$P(Y=High) = \frac{2}{3}$ $\quad$ $P(Y=Low) = 1/3$

Bayes Optimal Classifier $\hat{y} = \underset{y}{\operatorname{argmax}}\ P(y|x)$

$= \underset{y}{\operatorname{argmax}}\ p(x|y)\,p(y)$

Naive $\Rightarrow$ Cond. ind. assumption: $P(X_{1,2,\dots,d}) = P(X_1|Y) \cdot \ldots \cdot P(X_d|Y)$

for $i=1,2,3,4,5$ we calculated $P(X_i \mid Y=High)$ and $P(X_i \mid Y=Low)$, and $P(Y=High)$, $P(Y=Low)$

Doğukan Yıldırım 28364

(b)

for [Yes, No, Linear] $\xrightarrow{\text{by OHE}}$ [1, 0, 0, 0, 1]

$$\hat{y} = \underset{y}{\operatorname{argmax}}\ P(X_1, X_2, X_3, X_4, X_5 \mid Y=y), \quad y \in \{High, Low\}$$

For $y = High$, $P(X_1 = 1, X_2 = 0, X_3 = 0, X_4 = 0, X_5 = 1 \mid Y = High)$

$$= P(X_1 = 1 \mid Y = High)\, P(X_2 = 0 \mid Y = High)\, P(X_3 = 0 \mid Y = High)\, P(X_4 = 0 \mid Y = High)\, P(X_5 = 1 \mid Y = High)$$

$$= \frac{3}{4} \cdot \frac{3}{4} \cdot \frac{3}{4} \cdot \frac{1}{2} \cdot \frac{1}{4} = \frac{27}{512}$$

$$P(Y = High \mid X_1, X_2, X_3, X_4, X_5) = \frac{27}{512} \cdot P(Y = High)$$

$$= \frac{27}{512} \cdot \frac{2}{3}$$

$$= \frac{18}{512} = \frac{9}{256}$$

For $y = Low$, $P(X_1 = 1, X_2 = 0, X_3 = 0, X_4 = 0, X_5 = 1 \mid Y = Low)$

$$= P(X_1 = 1 \mid Y = Low)\, P(X_2 = 0 \mid Y = Low)\, P(X_3 = 0 \mid Y = Low)\, P(X_4 = 0 \mid Y = Low)\, P(X_5 = 1 \mid Y = Low)$$

$$= 0 \cdot 0 \cdot \frac{1}{2} \cdot 1 \cdot \frac{1}{2} = 0$$

$$P(Y = Low \mid X_1, X_2, X_3, X_4, X_5) = 0 \cdot \underbrace{\frac{1}{3}}_{P(Y = Low)}$$

$$= 0$$

$$\frac{9}{256} > 0$$

Therefore, $\hat{y} = High$.