Recap 2024-03-18 ①

- Matrix-Vector Notation for linear regression

In contrast to the scalar notation

$$y_i = b_0 + b_1 \cdot x_{1i} + e_i \quad \text{for animal } i$$

Data set on body weight and breast circumference

Intercept

$$\begin{aligned} y_1 &= \boxed{b_0} + b_1 \cdot x_{11} + e_1 \\ y_2 &= b_0 + b_1 \cdot x_{12} + e_2 \\ &\vdots \\ y_{10} &= b_0 + b_1 \cdot x_{110} + e_{10} \end{aligned} \qquad \begin{aligned} 471 &= b_0 + b_1 \cdot 176 + e_1 \\ 463 &= b_0 + b_1 \cdot 177 + e_2 \\ &\vdots \\ 541 &= b_0 + b_1 \cdot 184 + e_2 \end{aligned}$$

Matrix-Vector Notation: Define vectors y, b, e

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_{10} \end{bmatrix} = \begin{bmatrix} 471 \\ 463 \\ \vdots \\ 541 \end{bmatrix}; \quad b = \begin{bmatrix} b_0 \\ b_1 \end{bmatrix}; \quad e = \begin{bmatrix} e_1 \\ \vdots \\ e_{10} \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & 176 \\ 1 & 177 \\ \vdots & \vdots \\ 1 & 184 \end{bmatrix}; \quad \text{Model: } y = X \cdot b + e$$

Least Squares Estimate $\hat{b}$ for b

$$\hat{b} = (X^T X)^{-1} X^T y$$