

OHP Picture 1

Recap 2023-03-06

□ Regression :

- response / target variable $\rightarrow y$
- predictor variable $\rightarrow x$

$L = \sum_{i=1}^N e_i^2$; determine b_0 (intercept) and b_1 (slope) of regression line such that L is minimal

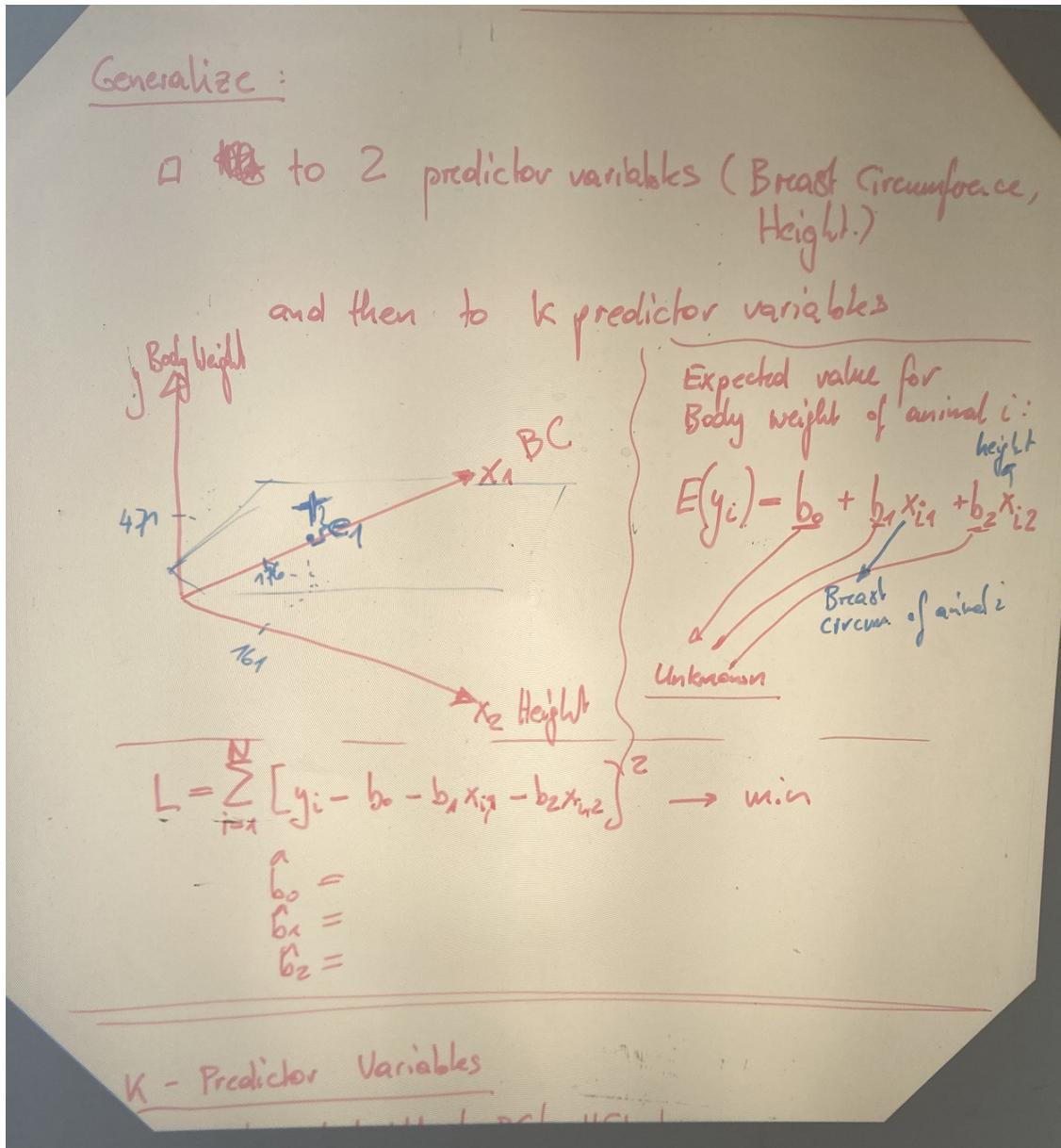
$$\hat{b}_0 = \bar{y} - \hat{b}_1 \bar{x}$$

$$\hat{b}_1 = \frac{(xy)_o - N\bar{x}\bar{y}}{(x^2)_o - N\bar{x}^2}$$

Generalize :

□ ~~to~~ to 2 predictor variables (Breast Circumference)

OHP Picture 2



OHP Picture 3

K - Predictor Variables

Ani	Body Weight y	BC x_1	HEI x_2	x_3 $x_4 \dots x_K$
-----	--------------------	---------------	----------------	-----------------------

N

$$L = \sum_{i=1}^N [y_i - b_0 - b_1 x_{i1} - b_2 x_{i2} - b_3 x_{i3} - \dots - b_K x_{ik}]^2 \rightarrow \min$$

$b_0 =$
 $b_1 =$
 \vdots
 $b_K =$

Simplified Notation : Matrix - Vector

Matrix : $X = \begin{bmatrix} x_{10} & x_{11} & x_{12} \\ x_{20} & x_{21} & x_{22} \\ \vdots & \vdots & \vdots \end{bmatrix}$

all = 1 Height
Intercept Breast circumference

Vectors : $y = \begin{bmatrix} y_1 \end{bmatrix}$: $I = \begin{bmatrix} b_0 \end{bmatrix} \rightarrow$ Intercept
Regression

OHP Picture 4

Simplified Notation : Matrix - Vector

• Matrix : $X = \begin{bmatrix} X_{10} & X_{11} & X_{12} \\ X_{20} & X_{21} & X_{22} \\ \vdots & \vdots & \vdots \end{bmatrix}$

↓ ↓ ↓
all = 1 Intercept Height
Body weight Breast circumference

Vectors : $y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}$; $b = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix}$

→ Intercept
→ Regression coefficient for BC
→ Regression coefficient for Height

Unknown

$$e = \begin{bmatrix} e_1 \\ \vdots \\ e_N \end{bmatrix}$$

$$\left\{ \begin{array}{l} E(y_1) = b_0 + b_1 x_{11} + b_2 x_{12} \\ E(y_2) = b_0 + b_1 x_{21} + b_2 x_{22} \end{array} \right.$$

$$E(y_i) = b_0 + b_1 x_{1i} + b_2 x_{2i} \rightarrow E(y_i)$$

$$\hookrightarrow E(y) = Xb$$

$$y = Xb + e \Rightarrow e = y - Xb$$

Now estimate b :

OHP Picture 5

How estimate \underline{b} :

□ Elements in vector \underline{b} are unknown
 \Rightarrow estimate from data using Least Squares

$$L = \underline{e}^T \underline{e} = [e_1 \ e_2 \ e_3 \dots e_N] \cdot \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{bmatrix} = e_1^2 + e_2^2 + \dots + e_N^2$$

$$L = \sum_{i=1}^N e_i^2 = e_1^2 + e_2^2 + \dots + e_N^2$$

$$\begin{aligned} L &= \underline{e}^T \underline{e} = (\underline{y} - \underline{X}\underline{b})^T \cdot (\underline{y} - \underline{X}\underline{b}) \\ &= (\underline{y}^T - (\underline{X}\underline{b})^T) \cdot (\underline{y} - \underline{X}\underline{b}) \\ &= (\underline{y}^T - \underline{b}^T \underline{X}^T) \cdot (\underline{y} - \underline{X}\underline{b}) \\ &= \cancel{\underline{y}^T \underline{y}} - \cancel{\underline{y}^T \underline{X}\underline{b}} - \underbrace{\underline{b}^T \underline{X}^T \underline{y}} + \underline{b}^T \underline{X}^T \underline{X}\underline{b} \\ &\quad (\underline{b}^T \underline{X}^T \underline{y})^T = \underline{y}^T \underline{X}\underline{b} \\ &= \underline{y}^T \underline{y} - 2\underline{y}^T \underline{X}\underline{b} + \underline{b}^T \underline{X}^T \underline{X}\underline{b} \end{aligned}$$

Least Squares : Find vector \underline{b} such that

OHP Picture 6

Least Squares : Find vector b such that

$\frac{\partial L}{\partial b_0}; \frac{\partial L}{\partial b_1}; \frac{\partial L}{\partial b_2}; \dots; \frac{\partial L}{\partial b_k}$ are all 0

Compute gradient of L with respect to b

$$\frac{\partial L}{\partial b} = \begin{bmatrix} \frac{\partial L}{\partial b_0} \\ \frac{\partial L}{\partial b_1} \\ \vdots \\ \frac{\partial L}{\partial b_k} \end{bmatrix} \quad \left\{ \begin{array}{l} L = y^T y - 2y^T K b + b^T K^T X b \\ \frac{\partial L}{\partial b} = 0 - 2y^T X + b^T X^T X \end{array} \right. \quad \left\{ \begin{array}{l} \frac{\partial x^n}{\partial x} = n x^{n-1} \\ = \text{---} \end{array} \right.$$

Find b such that $\frac{\partial L}{\partial b} = 0$

$$\Rightarrow -2y^T X + 2b^T X^T X = 0$$

$$b^T X^T X = y^T X$$

$$(X^T X) \hat{b} = X^T y \Rightarrow \text{Normal Equations}$$

Given that $(X^T X)$ can be inverted :

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

OHP Picture 7

Include also discrete variables as predictors

- Example :
 - Breed as an influence on body weight.
 - Breed is a discrete variable with fixed levels like
 $\{ \text{Angus, Simmental, Limousin} \}$

□ Data Set

Animal	Body Weight	Breed
1	471	1
2	.	
:		
10	541	2

- Expected body weight for animal i :
- $$E(y_i) = b_0 + b_1 \cdot x_i$$
- ↑ intercept ↓ regression coefficient → breed code

E.g. Animal 1: $E(y_1) = b_0 + b_1 \cdot 1$
 Animal 10: $E(y_{10}) = b_0 + b_1 \cdot 2$