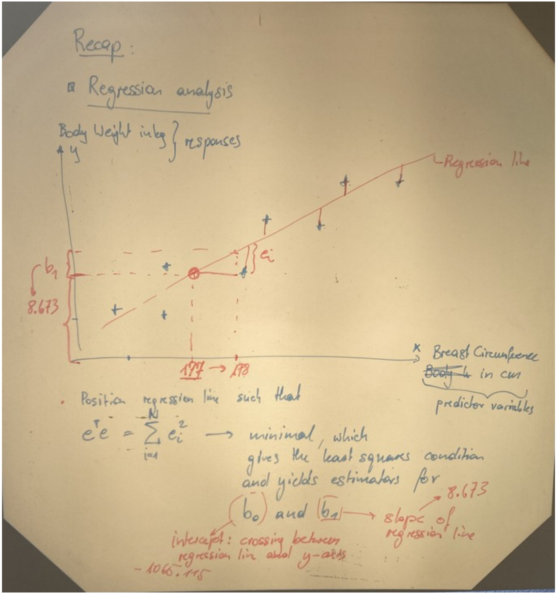
OHP Picture 1



OHP Picture 2

Extension of Linear Regression Model

- More than just one "x-variable"
- Example in Exercise 1:

Animal	x_1 BC (cm)	x_2 BCS (–)	x_3 HEI (cm)	y BW (kg)
1	176	5.0	161	471
2	⋮			
10				

- Are additional x-variable (BCS and HEI) bringing any new information for modelling BW?
- Extended model:
  - $E(y_i) = b_0 + b_1 \cdot x_1 + b_2 \cdot x_2 + b_3 \cdot x_3$
- Estimates for b_0, b_1, b_2 and b_3 are obtained from

$$\left.\frac{\partial e^Te}{\partial b_0}\,;\ \frac{\partial e^Te}{\partial b_1}\,;\ \frac{\partial e^Te}{\partial b_2}\,;\ \frac{\partial e^Te}{\partial b_3}\ \right\}\ \text{cumbersome} \ldots$$

OHP Picture 3

Matrix-Vector Notation for Linear Regression

- All x-variable are put into Matrix X :

$$\underline{X} = \begin{bmatrix} x_{10} & x_{11} & x_{12} \\ x_{20} & x_{21} & x_{22} \\ \vdots & \vdots & \vdots \\ x_{N0} & x_{N1} & x_{N2} \end{bmatrix} \quad N=10$$

column of all "1" BC HEI

$x_{10} = x_{20} = \ldots = x_{N0} = 1$

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}, \quad e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{bmatrix}, \quad b = \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} \begin{array}{l} \to \text{intercept} \\ \to \text{slope for BC} \\ \to \text{slope for HEI} \end{array}$$

responce vector BW

- Regression Model : $y = X \cdot b + e$ → the same for any number N and any number k of x-variables.

$$\begin{bmatrix} y_1 \end{bmatrix} = \begin{bmatrix} x_0 & x_{11} & x_{12} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix} \begin{bmatrix} e_1 \end{bmatrix}$$

$$y_1 = x_{10} \cdot b_0 + x_{11} \cdot b_1 + x_{12} \cdot b_2 + e_1$$

OHP Picture 4

Properties of Model in Matrix-vector notation

□ expected value of y:

$$E(y_1) = b_0 + b_1 x_{11} + b_2 x_{12}$$

$$E(y_2) = b_0 + b_1 x_{21} + b_2 x_{22}$$

$$\vdots$$

□ Define $E[y] = \begin{bmatrix} E(y_1) \\ E(y_2) \\ \vdots \\ E(y_N) \end{bmatrix} = \begin{bmatrix} b_0 + b_1 x_{11} + b_2 x_{12} \\ b_0 + b_1 x_{21} + b_2 x_{22} \\ \vdots \\ b_0 + b_1 x_{N1} + b_2 x_{N2} \end{bmatrix} = \begin{bmatrix} x_{10} & x_{11} & x_{12} \\ x_{20} & x_{21} & x_{22} \\ x_{N0} & x_{N1} & x_{N2} \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \\ b_2 \end{bmatrix}$

$$= X \cdot b$$

□ Verify: $E[y] = E[Xb + e] = E[Xb] + \underbrace{E[e]}_{=0} = E[Xb] + 0$

$$= E[Xb] = Xb$$

□ Error Terms: $E[e] = 0$; expected value

$\text{var}(e) = E[e - \underbrace{E(e)}_{0}][e - \underbrace{E(e)}_{=0}]^T = E[ee^T] = \sigma^2 I_N$

Identity matrix with dimension $N \times N$

OHP Picture 5

- Variance-Covariance Matrix $var(e)$

$$var(e) = \begin{bmatrix} var(e_1) & cov(e_1,e_2) & \cdots & cov(e_1,e_N) \\ cov(e_2,e_1) & var(e_2) & \cdots & cov(e_2,e_N) \\ \vdots & & \ddots & \\ \vdots & & & var(e_n) \end{bmatrix} = \sigma^2 \cdot I_N$$

$$= \begin{bmatrix} \sigma^2 & \emptyset & \emptyset & \cdots & \emptyset \\ \emptyset & \sigma^2 & & & \\ \vdots & & \ddots & & \\ \emptyset & & & & \sigma^2 \end{bmatrix}$$

$var(e_i) = \sigma^2 \Rightarrow$ where does it come from?

$\Downarrow$ estimated from data

$cov(e_i, e_j) = \emptyset \quad (i \neq j)$

- From the output of summary from lm() in R:
  "Residual standard error" is an estimate of σ

- $r =$ vector of residuals:

$$\hat{\sigma} = \sqrt{\frac{1}{N-2} r^T r} = \sqrt{\frac{1}{N-2} \sum_{i=1}^{N} r_i^2}$$

number of parameters $\Rightarrow b_0, b_1$

OHP Picture 6

Sum of Squared Residuals in Matrix-Vector Notation

- $e^T e = [y - Xb]^T \cdot [y - Xb] = y^T y -$

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