

Applied Statistical Methods - Solution 1

2022-02-25

Problem 1: Regression Analysis

The following dataset on body weight and on further observations on a number of animals is given.

Animal	Breast Circumference	Body Weight	BCS	HEI
1	176	471	5.0	161
2	177	463	4.2	121
3	178	481	4.9	157
4	179	470	3.0	165
5	179	496	6.8	136
6	180	491	4.9	123
7	181	518	4.4	163
8	182	511	4.4	149
9	183	510	3.5	143
10	184	541	4.7	130

The data can be read from https://charlotte-ngs.github.io/asmss2022/data/asm_bw_mult_reg.csv

Tasks

- Build a regression model of body weight on the other observations using the dataset given above.
- Set up the matrix \mathbf{X} and the vectors \mathbf{y} , \mathbf{b} and \mathbf{e} .
- Compute estimate for the regression coefficients in the model defined above.

Solution

First, the data has to be read

```
s_sol01_p01_path <- "https://charlotte-ngs.github.io/asmss2022/data/asm_bw_mult_reg.csv"
tbl_sol01_p01 <- readr::read_csv(file = s_sol01_p01_path)
```

The regression model can be stated as

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$$

where

$$X = \begin{bmatrix} 1 & 176 & 5 & 161 \\ 1 & 177 & 4.2 & 121 \\ 1 & 178 & 4.9 & 157 \\ 1 & 179 & 3 & 165 \\ 1 & 179 & 6.8 & 136 \\ 1 & 180 & 4.9 & 123 \\ 1 & 181 & 4.4 & 163 \\ 1 & 182 & 4.4 & 149 \\ 1 & 183 & 3.5 & 143 \\ 1 & 184 & 4.7 & 130 \end{bmatrix}, y = \begin{bmatrix} 471 \\ 463 \\ 481 \\ 470 \\ 496 \\ 491 \\ 518 \\ 511 \\ 510 \\ 541 \end{bmatrix}, e = \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ e_6 \\ e_7 \\ e_8 \\ e_9 \\ e_{10} \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix}$$

The solution for the estimates \hat{b} can be computed as

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

```
mat_xtx <- crossprod(mat_X)
mat_xty <- crossprod(mat_X, vec_y)
mat_sol_bhat <- solve(mat_xtx, mat_xty)
```

The solution vector is

$$\hat{b} = \begin{bmatrix} -1313.079 \\ 9.649 \\ 8.633 \\ 0.227 \end{bmatrix}$$

Problem 2

Solution