(1) Train a linear model $\log(Y) = \tilde{f}(X) = c_1 + c_2 X$ by

$$\tilde{f} = \textbf{lm}(\text{formula} = \log(Y) \sim X, \text{ data} = \dots).$$

(2) Predict Y by $Y = \exp(\tilde{f}(X))$, where

$$\tilde{f}(X) = \textbf{prediction}(\text{model} = \tilde{f}, \text{ newdata} = X).$$

(3) The coefficients $\beta_1 = \exp(c_1)$ and $\beta_2 = c_2$.

Linearization applies to many cases. For example, $Y = \beta_1 X_1 \cdot e^{\beta_2 X_2}$ can be converted to the linear model

$$\ln(Y/X_1) = \ln(\beta_1) + \beta_2 \cdot X_2.$$

To fit this model, one can use

$$\tilde{f} = \textbf{lm}(\text{formula} = \log(Y/X_1) \sim X_2, \text{ data} = \dots).$$

Note that the predicted value

$$\tilde{f}(X) = \textbf{prediction}(\text{model} = \tilde{f}, \text{ newdata} = X)$$

is just $\ln(Y/X_1)$. That is, $\tilde{f}(X) = \ln(Y/X_1)$. Therefore, the predicted Y value is

$$Y = X_1 \cdot \exp(\tilde{f}(X)).$$

If the model is $Y = X_1 \cdot e^{\beta X_2}$, then $\ln(Y/X_1) = \beta \cdot X_2$. and the model is trained by

$$\textbf{lm}(\text{formula} = \log(Y/X_1) \sim X_2 - 1, \text{ data} = \dots).$$

There are models which can't be linearized (called non-linear model). For example, the logistic model

$$Y = \frac{c}{1 + ae^{-bt}}.$$

It can also be solved, but with more complicated mathematical tools.

Example 1.5. Fit the data to the model $Y = \beta_1 \cdot \exp(\beta_2 X)$.

Project 6 - part 4 (3 points, due May 13)

Download data file "project6.csv" from Blackboard.

1. Fit the given data for model

$$X_5 = \beta_1 X_1 + \beta_2 X_2 + \beta_4 X_4.$$

2. Fit the given data for model

$$X_5 = \beta_1 X_1 e^{\beta_2 X_2 + \beta_4 X_4}.$$

3. Which model is better?