# Simple linear regression model — revisted using the maximum likelihood estimator

written by The Coué method on Functor Network
original link: https://functor.network/user/1751/entry/657

---

This is a revisit of the previous post. See https://functor.network/user/1751/entry/653.

## Goal

Write a mathematical model $y = \alpha + \beta x$ that describes the relationship between two variables $x$ and $y$.

## Setup

Given observations $(x_i, y_i)$, $i = 1, \ldots, n$, consider a model of the form

$$y_i = \alpha + \beta x_i + e_i$$

where $e_i$ are i.i.d. with $e_i \sim N(0, \sigma^2)$. The aim is to find estimates for $\alpha$, $\beta$, and $\sigma^2$.

### Likelihood function

The likelihood function $f_{x_1,x_2,\ldots,x_n,y_1,y_2,\ldots,y_n}(\alpha, \beta, \sigma^2)$ denoted $f$ is

$$f = \prod_{i=1}^{n} (2\pi\sigma^2)^{-\frac{1}{2}} \exp(-\frac{(y_i - \alpha - \beta x_i)^2}{2\sigma^2})$$

and the negative log likelihood function $L$ is

$$L = \sum_{i=1}^{n} \frac{1}{2} \log(2\pi\sigma^2) + \frac{(y_i - \alpha - \beta x_i)^2}{2\sigma^2}$$

or equivalently

$$L = \frac{n}{2} \log(2\pi\sigma^2) + \frac{1}{2\sigma^2} \sum_{i=1}^{n} {e_i}^2$$

where $e_i$ are the residuals.

**Solving for $\alpha$**

Computing $\frac{\partial L}{\partial \alpha}$ and setting it to zero yields

$$\frac{\partial}{\partial \alpha} \sum_{i=1}^{n} (y_i - \alpha - \beta x_i)^2 = 0$$

Per the work in the previous post, this results in the following estimate for $\alpha$.

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x} \tag{1}$$

**Solving for $\beta$**

Computing $\frac{\partial L}{\partial \beta}$ and setting it to zero yields

$$\frac{\partial}{\partial \beta} \sum_{i=1}^{n} (y_i - \alpha - \beta x_i)^2 = 0$$

Per the work in the previous post, this results in the following estimate for $\beta$.

$$\hat{\beta} = r_{xy} \frac{\sigma_y}{\sigma_x} \tag{2}$$

**Solving for $\sigma$**

Computing $\frac{\partial L}{\partial \sigma}$ and setting it to zero yields

$$\frac{\partial L}{\partial \sigma} = 0 \implies \frac{n}{\sigma} - \frac{1}{\sigma^3} \sum_{i=1}^{n} {e_i}^2 = 0$$

This gives the following estimate for $\beta$.

$$\sigma^2 = \frac{\sum_{i=1}^{n} {e_i}^2}{n} \tag{3}$$

The interpretation of equation (3) is that the sum of the squared residuals is an estimate of the variance of the error.

### The critical point is a local minimum of $L$ and therefore a local maximum of $f$

A computation shows that the Jacobian of $L$ at $(\hat{\alpha}, \hat{\beta}, \hat{\sigma})$ is

$$\begin{pmatrix} \frac{n}{\sigma^2} & \frac{n\bar{x}}{\sigma^2} & -\frac{2}{\sigma} \cdot \frac{\partial L}{\partial \alpha} \\ \frac{n\bar{x}}{\sigma^2} & \frac{\sum_{i=1}^n x_i^2}{\sigma^2} & -\frac{2}{\sigma} \cdot \frac{\partial L}{\partial \beta} \\ -\frac{2}{\sigma} \cdot \frac{\partial L}{\partial \alpha} & -\frac{2}{\sigma} \cdot \frac{\partial L}{\partial \beta} & -\frac{n}{\sigma^2} - \frac{3}{\sigma^4} \sum_{i=1}^n e_i^2 \end{pmatrix}$$

Using the critical point conditions, the Jacobian is

$$\begin{pmatrix} \frac{n}{\sigma^2} & \frac{n\bar{x}}{\sigma^2} & 0 \\ \frac{n\bar{x}}{\sigma^2} & \frac{\sum_{i=1}^n x_i^2}{\sigma^2} & 0 \\ 0 & 0 & \frac{2n}{\sigma^2} \end{pmatrix}$$

This matrix is positive definite because the (3,3) entry is positive and the minor $J_{3,3}$ is a positive definite $2 \times 2$ matrix, as shown in the previous post. Therefore the critical point is a local minimum of $L$ and a local maximum of $f$.

### Summing up

$$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$$

$$\hat{\beta} = r_{xy}\frac{\sigma_y}{\sigma_x}$$

$$\sigma^2 = \frac{\sum_{i=1}^n {e_i}^2}{n}$$

$$\hat{y} = \bar{y} + r_{xy}\frac{\sigma_y}{\sigma_x}(x - \bar{x})$$

### Comments

- The model is the same as the model in the previous post.

### Reference

Simple linear regression model, https://functor.network/user/1751/entry/653