Linear models

Juho Kopra

2023-02-09

1. Scatter plot and one linear predictor

Figure 1: scatterplot and a linear model with one predictor.

TODO: explain how we can fit a linear model to a scatter plot so that we minimise the sum of squared differences

Importantly, linear model is a statistical model which can be formulated as $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, and where $i = 1, \ldots, n$ and $\varepsilon_i \sim N(0, \sigma^2)$. The above definition we call general form of linear model as estimated values of parameters are not written down.

There the following terminology takes place

- Y_i is dependent
- β_0 is intercept
- β_1 is slope
- β_0 and β_1 are regression coefficients or model parameters
- x_i is the value of predictor aka independent aka covariate aka regressor

The fitting of linear model

We want to minimize the sum of squares of error terms ε_i^2, which can be written as $\varepsilon_i = Y_i - (\beta_0 + \beta_1 X_i)$. Note that error terms are in the same direction with Y-axis (and thus the same direction with Y_i).

We will not go into technical details of minimization here.

2. Properties of linear models and interpretation

Properties of linear models

To have an interpretation for β_0 and β_1 we note that:

Expected value of Y_i is if we consider x_i being known and fixed: $E(Y_i) = E(\beta_0 + \beta_1 x_i + \varepsilon_i) = E(\beta_0) + E(\beta_1 x_i) + E(\varepsilon_i) = E(\beta_0) + x_i E(\beta_1) + E(\varepsilon_i) = \beta_0 + \beta_1 x_i$, because x_i is constant and $E(\varepsilon_i) = 0$.

Shortly put, the above calculation gives $E(Y_i) = \beta_0 + \beta_1 x_i$, which is called a systematic part of linear model. If we replace the β-notation with their estimates, then we can use that equation to calculate fitted values: $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$ Fitted values are calculated for the same x_i values which were in the original data.

Prediction can be obtained by the same equation but using any value.

Interpretation

An interpretation of β_1 becomes available as

$E(Y_i|x_i) - E(Y_i|x_i + 1) = \beta_0 + \beta_1 x_i - (\beta_0 + \beta_1(x_i + 1)) = \beta_1 x_i - \beta_1(x_i + 1) = \beta_1 x_i - \beta_1 x_i + \beta_1 \cdot 1 = \beta_1$

Thus, if e.g. $\beta_1 = 5.3$ then the expected value of Y_i increases by 5.3 if x_i becomes increased with 1 unit.

An interpretation of intercept term β_0 becomes available as

3. Categorical predictors

4. Basics of Matrix algebra

5. Linear models formulated with Matrix algebra

4. Linear hypothesis testing with L-matrices

5. Contrast matrices

6. Confounding