FINC 305 LN2 - Linear Regression

Linxian Huang

December 27, 2024

In both Economics and Finance studies, we always begins with the following premise: y and x are two variables, representing some data we retrieved from real world, and we are interested in "how y varies with changes in x", such as "how return of Apple's stock change when the market portfolio return changes." Based on this consideration, y is called the **dependent variable**, the **independent variable**, **response variable**, or **predicted variable**. x is called **independent variable**, the **explanatory variable**, the **control variable**, the **predictor variable** or the **regressor**. Also, we can easiliy imagine that it has more than one factor that can significantly influence Apple's stock return. In this lecture, we will go further to review and study simple and multiple regresssions, discussing potential applications, and how to inprove the explanatory and predictive power of linear models you build.

1 Simple and Multiple Linear Regression

Linear regression model tries to explain a *dependent variable* (y) and one or more independent variables x_i. Just like how I interpreted previously, it has lots of terminology get used. A generic form of linear regression is:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \epsilon$$

where:

- y is the dependent variable,
- β_0 is the intercept,
- $\beta_1, \beta_2, \ldots, \beta_p$ are the coefficients,

- $x_1, x_2, \ldots, x_p$ are the independent variables,
- ϵ is the error term.

By converting this regression model to a matrix form, it can be interpreted as:

$$Y = X\beta + \epsilon$$

where:

- Y is an $n \times 1$ vector of dependent variables,
- X is an $n \times k$ matrix of independent variables,
- β is a $k \times 1$ vector of coefficients,
- ϵ is an $n \times 1$ vector of errors.

For a single observation i, the model becomes:

$$y_i = x_i^\top \beta + \epsilon_i$$

where:

- y_i is a scalar dependent variable,
- x_i is a $k \times 1$ vector of independent variables,
- β is a $k \times 1$ vector of coefficients,
- ϵ_i is a scalar error term.

1.1 Assumption