Loss Functions

Tyler J. Brough

1/19/2022

Another Look at OLS

Suppose we are working with a model of the form

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

- we have n data points (x_i, y_i), for $i = 1, \ldots, n$
- we seek to estimate estimate α and β.

The Loss Functions

Our parameters are the following: $\theta = \{\alpha, \beta, \sigma\}$

Applying the least squares criterion implies the following loss function:

$$\hat{\theta} = \arg\min_{\theta} \left[\sum_{i=1}^{n} (y_i - \beta x_i - \alpha)^2 \right]$$

The Solution

A necessary condition for optimality is that the two partial derivatives $\partial S/\partial \alpha$ and $\partial S/\partial \beta$ equal zero, giving us the following equations:

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

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

Where $S = \sum_{i=1}^{n}(y_i - \beta x_i - \alpha)^2$

The Solution Continued

We can rewrite the above equations in the form of the **normal equations** as follows:

$$\beta \sum_{i=1}^{n} x_i^2 + \alpha \sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i$$

$$\beta \sum_{i=1}^{n} x_i + n\alpha = \sum_{i=1}^{n} y_i$$

The Solution Continued

The normal equations can be solved for α and β to yield the following:

$$\hat{\beta} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{n\sum x_i^2 - (\sum x_i)^2}, \quad \text{the } \textbf{slope}$$

$$\hat{\alpha} = \frac{\sum x_i^2 \sum y_i - \sum x_i y_i \sum x_i}{n\sum x_i^2 - (\sum x_i)^2}, \quad \text{the } \textbf{intercept}$$

- Thus the solution is available in closed-form
- Given our data $(x_i, y_i)_{i=1}^n$ we can simply plug-and-chug to get point estimates

OLS Revisited

Thus the least squares approach uses the following loss function:

$$L(\theta) = S(\theta) = \arg\min_{\theta} \left[\sum_{i=1}^{n} (y_i - \beta x_i - \alpha)^2 \right]$$

Is this the *right* loss function?