ECON 61001: Review of OLS in simple linear regression

Alastair R. Hall

The University of Manchester

Outline of this revision session

- Origins of OLS
- Simple linear regression model
- Intuition behind OLS
- Formal definition of OLS
- Other approaches and advantages of OLS

Resources: any introductory econometrics text such as Wooldridge (2019) *Introductory Econometrics*.

Origins of OLS

The origin of Ordinary Least Squares (OLS) is shrouded in controversy.

The method is published by

- Adrien Legendre, a French mathematician, in 1805.
- Robert Adrain, an American in 1808.
- Carl Friedrich Gauss, a German, in 1809
    - but Gauss claimed he had been using the method since 1795.

Stigler (1981): "The method of least squares is the automobile of modern statistical analysis" but who was the "Henry Ford of statistics"?

Origins of OLS

Gauss was a very eminent mathematician whose key contributions include:

- the Normal (or “Gaussian”) distribution.
- development of Least Squares theory
- Gaussian elimination

Stigler (1981) evaluates Gauss's claim using historical records and concludes

> *Just as the automobile was not the product of one man of genius, so too the method of least squares is due to [...] at least two independent discoverers. Gauss may well have been the first of these, but he was no Henry Ford of statistics. If there was any single scientist who first put the method within the reach of the common man, it was Legendre.(Stigler, 1981,p.472)*

Suppose we are interesting in modeling the relationship between y, the annual salary of the CEO of a firm i, and x_i the average return on equity for the CEO's firm for the previous 3 years.

Assume simple linear regression model

$$y_i = \beta_{0,1} + \beta_{0,2}x_i + u_i = x_i'\beta_0 + u_i$$

where

- u_i is the unobserved error term;
- $x_i' = (1, x_i)$;
- $\beta_0 = \begin{bmatrix} \beta_{0,1} \\ \beta_{0,2} \end{bmatrix}$ are unknown (regression) parameters.

Under assumptions discussed in the lecture, we have

$$E[y_i \,|\, x_i] \;=\; \beta_{0,1} \,+\, \beta_{0,2} x_i$$

So $E[y_i \,|\, x_i]$ is linear function of x_i but weights (the regression parameters) are unknown.

So collect sample on $\{y_i, x_i\}$ and use this to estimate β_0.

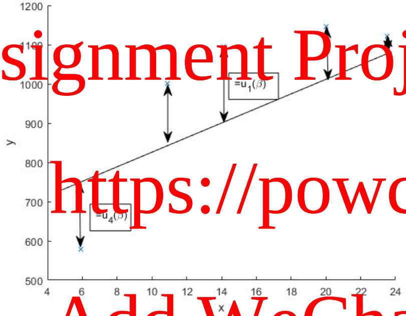
Suppose our data consists of the following five observations on (y_i, x_i):

$\{(1095, 14.1), (1001, 10.9), (1122, 23.5), (578, 5.9), (1145, 20.0)\}$.

No single line passes through all the points.

Choose line that comes "closest" to fitting the scatter plot $\rightarrow$ issue of how to measure distance of actual y from value predicted by line, $\beta_1 + \beta_2 x$.

Define $u_i(\beta) = y_i - \beta_1 - \beta_2 x_i$



Measure of distance from line must be non-negative: in OLS we measure this distance by $\{u_i(\beta)\}^2$

OLS estimator of (β_1, β_2) is value that minimizes $\sum_{i=1}^{5}\{u_i(\beta)\}^2$.

In this example the OLS line is:

$$\hat{y}_i = 573.63 + 27.86 x_i$$

Notice that $\{u_i(\beta)\}^2$ is just one possible measure of distance.

For example, could also use $|u_i(\beta)|$

If choose (β_1, β_2) to minimize $\sum_{i=1}^{5} |u_i(\beta)| \rightarrow$ Least Absolute Deviation (LAD) regression

LAD actually proposed in 1757 by Roger Boscovich but OLS became the "automobile of modern statistical analysis" because:

- calculus of OLS is far easier
- OLS can be shown to have some desirable statistical properties (see lectures).