The linear regression model with multiple regressors, and the standard deviation estimator of the OLS estimator

Econometrics for minor Finance, Lecture 5

Tunga Kantarcı, Fall 2025

Estimating the standard deviation of the OLS estimator: SD estimator

The standard error estimator of OLS estimator $\hat{\beta}_j$ in the linear regression model with multiple predictors is given by

$$\text{SEE}\left[\hat{\beta}_j \,\middle|\, x\right] = \sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2 (1 - R_j^2)}}$$

if errors are homoskedastic. Here

$$R_j^2$$

is the R^2 from a regression of x_j on all other regressors.

Estimating the standard deviation of the OLS estimator: SD estimator: Determinants

$$\text{SEE}\left[\hat{\beta}_j \,\middle|\, x\right] = \sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2 \left(1 - R_j^2\right)}}$$

The expression shows that the SEE of the OLS estimator is

i. higher if the estimated variance of the regression error $\hat{\sigma}^2$ is higher,

ii. lower if the sample size n is larger,

iiii. lower if the sample variation in the predictor $x_i - \bar{x}$ is larger,

iv. larger if correlation between x_j and other regressors is larger meaning if R_j^2 is larger.

Estimating the standard deviation of the OLS estimator: SD estimator: Determinants: Multicollinearity

It is worth to be more specific about the last determinant.

Estimating the standard deviation of the OLS estimator: SD estimator: Determinants: Multicollinearity

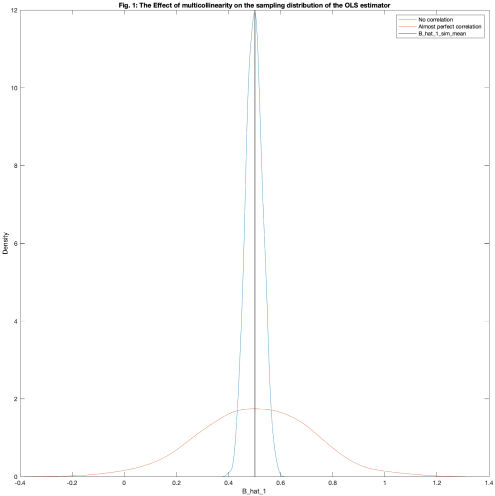
R_j^2 will always be positive because explanatory variables will always be correlated to some extent, even spuriously. So some correlation between explanatory variables is not a problem.

What is a problem is if they are strongly correlated, a phenomenon known as multicollinearity. When explanatory variables are strongly correlated, they share a significant amount of common information, which makes it challenging to estimate their individual effects accurately. This leads to inflated S.E.s for the estimated regression coefficients, resulting in less reliable estimates.

Estimating the standard deviation of the OLS estimator: SD estimator: Determinants: Multicollinearity

As we are talking about the variance of the OLS estimator, we can check its sampling distribution under different levels of correlation between explanatory variables. Let's check extreme cases of perfect correlation and no correlation.

Estimating the standard deviation of the OLS estimator: SD estimator: Determinants: Multicollinearity



Estimating the standard deviation of the OLS estimator: SD estimator: Determinants: Multicollinearity

S.D. of the sampling distribution of an OLS estimator under different correlation levels:

Estimating the standard deviation of the OLS estimator: SD estimator: Determinants: Multicollinearity