

Lecture 18: OLS estimates too large on average;
practical consequences; detection by $|X'X|$
Econometrics ■ Unit II: Multicollinearity and heteroskedasticity

Kapil Kumar

Department of Statistics
Chaudhary Charan Singh University, Meerut

B.Sc. (Hons) Statistics, Semester V
Academic Session 2026–2027

Previous lecture

Lecture 17: Implications: perfect and near multicollinearity; the variance inflation factor.

Today

- Implications (consequences) of multicollinearity
 - OLS estimates are too large on average
 - Summary of practical consequences
- Tools for handling the problem: diagnostics
 - Determinant of $X'X$

Let $L^2 = (b - \beta)'(b - \beta) = \sum_j (b_j - \beta_j)^2$ be the squared distance between b and β .

Step 1. Since $E(b_j) = \beta_j$,

$$E(L^2) = \sum_j E(b_j - \beta_j)^2 = \sum_j \text{Var}(b_j) = \sigma^2 \text{tr}(X'X)^{-1}.$$

Step 2. If $\lambda_1, \dots, \lambda_k$ are the eigenvalues of $X'X$, the eigenvalues of $(X'X)^{-1}$ are $1/\lambda_1, \dots, 1/\lambda_k$. The trace is the sum of the eigenvalues, so

$$E(L^2) = \sigma^2 \sum_{j=1}^k \frac{1}{\lambda_j}.$$

Step 3. Expanding $L^2 = b'b - 2\beta'b + \beta'\beta$ and taking expectations,

$$E(L^2) = E(b'b) - 2\beta'\beta + \beta'\beta = E(b'b) - \beta'\beta.$$

Therefore

$$E(b'b) = \beta'\beta + \sigma^2 \sum_{j=1}^k \frac{1}{\lambda_j}.$$

- Under multicollinearity at least one λ_j is small, so $1/\lambda_j$ is large.
- Then $b'b$ is on average much larger than $\beta'\beta$: the OLS estimates tend to be too large in absolute value.

In the presence of harmful (near) multicollinearity:

1. **OLS remains unbiased and BLUE**, because no OLS assumption is violated. But its variances are very large, so the estimates are imprecise. “Best” only means best among linear unbiased estimators; the best may still be poor.
2. **Coefficients may appear insignificant.** To test $H_0 : \beta_1 = 0$ we use $t_0 = b_1 / \sqrt{\widehat{\text{Var}}(b_1)}$. When $\text{Var}(b_1)$ is large, t_0 is small and H_0 is accepted more often. Important variables may then be wrongly dropped.
3. **Confidence intervals become wide.** The interval $b_1 \pm t_{\alpha/2, n-k-1} \sqrt{\widehat{\text{Var}}(b_1)}$ widens as $\text{Var}(b_1)$ grows.

4. **Estimates become unstable.** Adding or deleting a few observations can change the estimates greatly, even their signs.
5. **High R^2 with few significant t ratios** is a common sign.
6. **Prediction may still be good.** Multicollinearity damages the estimation of individual coefficients, not necessarily the fitted values. Predictions at values of the regressors that follow the same pattern as the sample can be satisfactory.

Exam tip

A common mistake is to say that OLS is not BLUE under multicollinearity. OLS remains unbiased and BLUE; the problem is that its variances become very large.

Detection of multicollinearity has three aspects:

- (i) determining its presence,
- (ii) determining its severity, and
- (iii) determining its form, that is, which variables are involved.

No single diagnostic is best; several are used together.

When multicollinearity increases, $X'X$ approaches singularity, so $|X'X| \rightarrow 0$. Exact multicollinearity gives $|X'X| = 0$. Thus a small $|X'X|$ suggests multicollinearity.

Limitations.

- (i) For unscaled data $|X'X|$ is not bounded: $0 \leq |X'X| < \infty$.
- (ii) It depends on the spread of the variables. For two centred (but not scaled) variables,

$$|X'X| = \begin{vmatrix} \sum x_{1i}^2 & \sum x_{1i}x_{2i} \\ \sum x_{1i}x_{2i} & \sum x_{2i}^2 \end{vmatrix} = \sum x_{1i}^2 \sum x_{2i}^2 - \left(\sum x_{1i}x_{2i} \right)^2.$$

(ii) (cont.) Since $\sum x_{1i}x_{2i} = r_{12}\sqrt{\sum x_{1i}^2 \sum x_{2i}^2}$, this equals

$$|X'X| = \sum x_{1i}^2 \sum x_{2i}^2 (1 - r_{12}^2).$$

So $|X'X|$ can be small simply because the variables have little variability, even when r_{12} is small.

(iii) It does not tell which variables cause the problem.

Practice

Re-work the derivations and examples of this lecture without looking at the slides.

Next lecture

Lecture 19: Diagnostics: correlation matrix, $|R|$, partial regressions and VIFs.