Gaus Markov Theorem

EC 320: Introduction to Econometrics

Winter 2022

Classical Assumptions

Classical Assumptions of OLS

- **A1. Linearity:** The population relationship is **linear in parameters** with an additive error term.
- **A2. Sample Variation:** There is variation in X.
- **A3. Exogeneity:** The X variable is **exogenous** (*i.e.*, $\mathbb{E}(u|X) = 0$)[†]
- **A4. Homoskedasticity:** The error term has the same variance for each value of the independent variable (*i.e.*, $\mathrm{Var}(u|X) = \sigma^2$).
- **A5. Non-autocorrelation:** The values of error terms are independent from one another (*i.e.*, $E[u_i u_j] = 0, \forall i \text{ s.t. } i \neq j$).[†]
- **A6. Normality:** The population error term is normally distributed with mean zero and variance σ^2 (*i.e.*, $u \sim N(0, \sigma^2)$)

† **Random Sampling:** Notice up until now, our underlying data type used for analysis was cross-sectional data. Under random sampling, this yields u_i and u_j independent for any two observations i and j, which is what non-autocorrelation assumption implies. However here I explicitly state **non-autocorrelation** to 1) generalize this case to account for a case we use time series data, 2) to be consistent with the notations from the textbook. Also it could be shown that the errors for different observations are independent conditional on the explanatory variables in the cross-sectional sample under random sampling.

Gauss-Markov Theorem

Gauss-Markov Theorem

OLS is the **Best Linear Unbiased Estimator** under assumptions A1-A5:

- **A1. Linearity:** The population relationship is **linear in parameters** with an additive error term.
- **A2. Sample Variation:** There is variation in X.
- **A3. Exogeneity:** The X variable is **exogenous** (*i.e.*, $\mathbb{E}(u|X) = 0$).
- **A4. Homoskedasticity:** The error term has the same variance for each value of the independent variable (*i.e.*, $\mathbf{Var}(u|X) = \sigma^2$).
- **A5. Non-Autocorrelation:** Any pair of error terms are drawn independently of each other (*i.e.*, $\mathbf{E}(u_i u_j) = 0 \; \forall \; i \; \text{s.t.} \; i \neq j$)

Gauss-Markov Theorem

OLS is the **Best Linear Unbiased Estimator (BLUE)**

- **Best:** most efficient (all other estimators will have a greater variance)
- **Linear:** as apposed to non-linear population model
- **Unbiased:** the expected value of the estimators are the population parameters