The Classical Linear Regression model

- Some conclusions

Textbooks as Catalogs

There are no super estimators – some estimators have desirable properties under some circumstance. Textbook can be characterized as a catelog of which estimators are more desirable in what estimating situations.

The catalog is centred around the CLR model, classical linear regression model. The OLS estimator is the optimal estimator under five assumptions.

1. linear relationship
   a. wrong regressors
   b. nonlinearity
   c. changing parameters
2. zero value for the unconditional expected value of the error term – this leads to biased intercept term.
3. spherical variance-covariance matrix for the error terms.
   a. heteroskedasticity
   b. autocorrelation in the error term
4. exogeneity for the regressors
5. no perfect multicollinearity

The assumptions of the CLR model

1. The model is correctly specified.

$$y_t = \beta_0 + \beta_t x_t + u_t, t = 1, \dots, n, u_t \sim IID(0, \sigma_0^2)$$

   Or in a matrix notation:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{u}, \boldsymbol{u} \sim \boldsymbol{IID}(0, \sigma_0^2 \boldsymbol{I})$$

2. The expected value of the error term is zero:

$$E(u_t) for\ all\ t, \text{ or } E(\boldsymbol{u}) = 0$$

3. Spherical variance-covariance matrix, e.g., no serial correlation and Heteroskedasticity:

$$E\left(u_t u_{t-j}\right) = 0, \text{ when } j \geq 1$$

$$E(u_t u_{t-j}) = \sigma_0^2, \text{ when } j = 0$$

Or

$$E(\boldsymbol{u}\boldsymbol{u}') = \sigma_0^2 \boldsymbol{I}$$

4. Exogeneity or fixed regressor in repeated samples:

When we assume 'fixed regressor in repeated samples', we assume that we can redraw the observations with the same explanatory variable values (e.g., we are interested in the impact of each state's economy on their residents' salary. Thus, every time we redraw the sample from a new cohort of people in those states, we keep the variable for the states' economy as fixed. In this example we exhaust the population of the regressor). This is an assumption stronger than independence between the regressor and the error terms, which is stronger than exogeneity, which is stronger than predeterminedness. Note that, predeterminedness indicates that:

$$E(u_t|x_t) = E(u_t)$$

We assume predeterminedness when we use the OLS / MM estimator by definition, but the estimator will have properties such as unbiasedness only when we at least assume exogeneity, e.g.,

$$E(u_t|x_{t-j}) = E(u_t), \text{ where } j = 0, 1, \ldots$$

Or

$$E(\boldsymbol{u}|\boldsymbol{X}) = E(\boldsymbol{X})$$

5. No perfect multicollinearity

OLS has the following properties:

1. Low computational cost
2. Least square by definition
3. Highest R^2
4. BLUE, when the assumptions above are valid.
5. MSE of the parameters. OLS is unbiased, but some other biased estimator may yield smaller MSE of the parameters.
6. OLS is unbiased in finite samples and asymptotically. It is consistent. It is also efficient asymptotically efficient in the CNLR model.

7. ML. we cannot calculate the maximum likelihood based on the assumptions of the CLR model as it does not specify the distribution of the error terms. However, if we can assume the error terms are normally distributed, then $\boldsymbol{\beta}^{MLE}$ is equivalent to $\boldsymbol{\beta}^{OLS}$

The Ballentine is helpful to understand the OLS estimator (e.g., p45, Kennedy, 2008).

A linear restriction on the parameters of the regression model can be incorporated into a regression buy eliminating one coefficient from that equation and running the resulting regression using transformed variables.