Autocorrelation

Class 4

Wonmun Shin

(wonmun.shin@sejong.ac.kr)

Department of Economics, Sejong University

* This lecture note is written based on Professor Chang Sik Kim's lecture note.

What is Autocorrelation?

What is Autocorrelation?

- **Autocorrelation**: Auto + Correlation (also known as **serial correlation**)
- In classical assumptions, regression errors are assumed to be uncorrelated, that is,

$$Cov\left(e_i, e_j\right) = 0 \quad \text{for } i \neq j$$

- Here, we relax the above assumption by allowing for correlation in the error terms.

What is Autocorrelation? [cont'd]

- When you have time series data, where the observations follow a natural ordering through time, there is always a possibility that successive error will be correlated each other.
- Consider a simple macro-econometric model with macroeconomic variables, such as income and consumption.
  - These variables can be related with the variables in the previous periods which can cause the autocorrelation.
- Sources of autocorrelation
  - Inertia (or slugishness)
  - Specification error: excluded variables, incorrect functional form
  - Cobweb phenomenon
  - Manipulation of data

What is Autocorrelation? [cont'd]

- **Autocorrelation**: Consider the following regression

$$Y_t = \beta_1 + \beta_2 X_t + e_t \quad \text{for } t = 1, 2, \cdots, n$$

then we have

$$Cov\left(e_t, e_s\right) \neq 0 \quad \text{for } t \neq s$$

- As we change our notation from i to t for the subscript, the autocorrelation is mostly encountered when using time series data.

Consequences of Autocorrelation

Consequences of Autocorrelation

$$Y_t = \beta_1 + \beta_2 X_t + e_t$$

- **Question 1** Can we obtain the OLS estimators under autocorrelation?
  - Note that we did not use (A4) to construct the OLS estimators.
  - $\therefore$ We can earn $\hat{\beta}_1$ and $\hat{\beta}_2$ through the ordinary least square method.

$$\hat{\beta}_2 = \frac{\sum x_t y_t}{\sum x_t^2}$$

Consequences of Autocorrelation [cont'd]

- **Question 2** Then, does still the properties of OLS estimator stand?
  - Unbiasedness

$$\hat{\beta}_2 - \beta_2 = \sum \omega_t e_t$$
$$\rightarrow E\left(\hat{\beta}_2 - \beta_2\right) = E\left(\sum \omega_t e_t\right) = \sum \omega_t E\left(e_t\right) = 0$$

    - So, $\hat{\beta}_2$ is a still unbiased (and linear) estimator.
  - Variance of OLS estimator

$$\begin{aligned} Var\left(\hat{\beta}_2\right) &= Var\left(\hat{\beta}_2 - \beta_2\right) = Var\left(\sum \omega_t e_t\right) \\ &= \omega_1^2 Var\left(e_1\right) + \cdots + \omega_n^2 Var\left(e_n\right) \\ &\quad + \omega_1 \omega_2 Cov\left(e_1, e_2\right) + \cdots + \omega_n \omega_{n-1} Cov\left(e_n, e_{n-1}\right) \\ &= \sigma^2 \sum \omega_t^2 + 2 \sum_{t \neq s} \omega_t \omega_s Cov\left(e_t, e_s\right) \\ &= \frac{\sigma^2}{\sum x_t^2} + 2 \sum_{t \neq s} \omega_t \omega_s Cov\left(e_t, e_s\right) \end{aligned}$$

    - The usual formula for the variance of the OLS estimator is incorrect.

Consequences of Autocorrelation [cont'd]

- Consistency
  - Recall that the sufficient condition of consistency is $MSE \to 0$ as $n \to \infty$.
  - $Bias\left(\hat{\beta}_2\right) = 0$ and $Var\left(\hat{\beta}_2\right) \to 0$ as $n \to \infty$
  - $\therefore$ $\hat{\beta}_2$ is a still consistent estimator.
- **Question 3** We have different formula for the variance of the OLS estimator. Then, the OLS estimator is still BLUE?
  - Gauss-Markov theorem does not apply any more, so OLS estimators are *no longer BLUE*.
  - There exists another linear unbiased estimator of β_2 which has a smaller variance than $\hat{\beta}_2$ when there is no autocorrelation (*Generalized Least Squares (GLS)* estimator).
  - Nevertheless, it doesn't mean we cannot use the OLS estimator any more.
    - But we should sacrifice the accuracy of estimator owing to the larger variance.
    - To make it worse, if we ignore the existence of autocorrelation and use the usual OLS estimator, the inference would be misleading!

AR(1) Process Error

First-Order Autoregressive Process: AR(1)

- Since the assumption $Cov\left(e_t, e_s\right) = 0$ is no longer valid, we need to replace this assumption with an alternative to describe the correlation among error terms.

- There are lots of models that can be used to represent the correlated errors.

- We will limit our attention to a model to a AR(1) model which is one of the most commonly-used simple method of autocorrelation modelling.

First-Order Autoregressive Process: AR(1) [cont'd]

- **AR(1) model**: In this model, e_t depends on its lagged value e_{t-1} plus another random component, that is

$$e_t = \rho e_{t-1} + v_t, \quad \text{for } t = 1, \cdots, n$$

  - ρ is a parameter that determines the correlation properties of e_t, and indicates the magnitude of the previous period's effect on e_t.
  - v_t is uncorrelated random variable with constant variance.

$$E(v_t) = 0, \ \ Var(v_t) = \sigma_v^2, \ \ Cov(v_t, v_s) = 0 \ \text{ for } t \neq s$$

  - ρe_{t-1} is a carryover from the previous period due to the inertia in economic system.
  - v_t is a new shock to the level of the economic variable.

First-Order Autoregressive Process: AR(1) [cont'd]

- Properties of AR(1) process error
  - Assumption: ρ is less than one in absolute value, that is

$$-1 < \rho < 1$$

  - Zero mean: $E(e_t) = 0$
  - Homoskedastic variance: $Var(e_t) = \sigma_e^2 = \frac{\sigma_v^2}{1-\rho^2}$
    - Note This variance does not change over time.
  - Covariance: $Cov(e_t, e_{t-k}) = \rho^k \sigma_e^2$ for $k \geq 0$
    - Correlation coefficient: $Corr(e_t, e_{t-k}) = \frac{Cov(e_t, e_{t-s})}{\sqrt{Var(e_t)} \cdot \sqrt{Var(e_t)}} = \rho^k$
    - Note $Corr(e_t, e_{t-1}) = \rho$

Detecting Autocorrelation

1. Durbin-Watson Test (DW Test)

- Assumptions
  1. Error term e_t follows AR(1) process. ($\rightarrow$ **Test for AR(1) error**)
  2. e_t follows normal distribution. ($\leftarrow$ small sample test)
  3. The regression model includes the intercept term.
  4. The regression model does not include the lagged values of Y_t.
- Test

$$\begin{cases} H_0 : \rho = 0 & (\rightarrow e_t = v_t, \text{ No autocorrelation}) \\ H_1 : \rho \neq 0 \end{cases}$$

- Durbin-Watson statistic (d statistic)

$$d = \frac{\sum_{t=2}^{n} \left(\hat{e}_t - \hat{e}_{t-1}\right)^2}{\sum_{t=1}^{n} \hat{e}_t^2}$$

where $\hat{e}_t$ are residuals from OLS regression ($i.e.$ $\hat{e}_t = Y_t - \hat{\beta}_1 - \hat{\beta}_2 X_t$).

Durbin-Watson Test [cont'd]

- It can be shown that approximately we have

$$d \approx 2\left(1 - \hat{\rho}\right)$$

where

$$\hat{\rho} = \frac{\sum_{t=2}^{n} \hat{e}_t \hat{e}_{t-1}}{\sum_{t=1}^{n} \hat{e}_t^2}$$

- It follows that

$$\begin{array}{rll} \hat{\rho} \approx 0 & \text{No autocorrelation} & d \approx 2 \\ \hat{\rho} \approx -1 & \text{Negative correlation} & d \approx 4 \\ \hat{\rho} \approx 1 & \text{Positive correlation} & d \approx 0 \end{array}$$

- From the Durbin-Watson table (we don't know the distribution of d, but can know its upper bound and lower bound), we test the null of no autocorrelation.

Durbin-Watson Test [cont'd]

- Caveats
  - Test for AR (1) error: If true error does not follow AR(1)?
  - Normality assumption
  - Inconclusive region may be large ($\because$ we don't know the distribution of d)
  - No allowing lagged dependent variable as regressors

Durbin's h Test

2. Durbin's h test

- Asymptotic test $\rightarrow$ we don't need distributional assumption.
- It allows one lagged dependent variable as a regressor.
- Test

$$\begin{cases} H_0 : \rho = 0 \\ H_1 : \rho \neq 0 \end{cases}$$

- Test statistic (h statistic) converges to standard normal, so we can test it.
- Caveats
  - Still, h test is only for AR(1) error.
  - Only one lagged dependent variable is allowed as a regressor
- Note h test can be understood as a *"stepping stone"* to BG test!

3. Breusch-Godfrey LM Test

- Breusch and Godfrey have developed a general test of autocorrelation.
  - It allows for the lagged values of the dependent variables.
  - It allows higher-order autoregressive schemes.
- Test

$$\begin{cases} H_0 : \ \textit{No autocorrelation} \\ H_1 : \ e_t \sim AR(p) \end{cases}$$

- Auxiliary regression: $\hat{e}_t = \alpha_1 + \alpha_2 X_t + \rho_1 \hat{e}_{t-1} + \cdots + \rho_p \hat{e}_{t-p} + u_t$
- We can compute $R^2 \quad \to n \cdot R^2 \overset{asy}{\sim} \chi^2_p$
- If we reject the null ($n \cdot R^2 >$ critical value), $\exists$ autocorrelation.
- Caveats Choice of p (the length of the lag) is arbitrary.

Others runs test, Box-Pierce test, Ljung-Box test, $\cdots\cdots$

Solutions for Autocorrelation

1. Newey-West HAC Estimator

- Given that the conventional lest squares are incorrect under the correlation, the **Newey-West HAC estimator** gives you a consistent estimator for the variance of OLS estimator.
- **HAC** (Heteroskeadsticity and Autocorrelation Consistent) estimator: Newey-West estimator corrects autocorrelation as well as heteroskeadsticity.
- However, OLS estimator $\hat{\beta}_2$ is still inefficient (since Gauss-Markov theorem no longer holds).

Solutions for Autocorrelation: GLS

2. Generalized Least Squares (GLS)

- Like in the case of heteroskedasticity, we transform the model so that the transformed model satisfies the classical assumptions.
- Then, GLS which takes the autocorrelation into account explicitly obtains the minimum variance within the class of linear unbiased estimators.
- It is so-called *Cochrane-Orcutt transformation*.

- Assume that error term follows AR(1) and we know its persistence ρ.
- Consider the simple regression model as

$$Y_t = \beta_1 + \beta_2 X_t + e_t \quad \text{for } t = 1, 2, \cdots, n \tag{1}$$

where the regression error follows

$$e_t = \rho e_{t-1} + v_t, \quad \text{for } t = 1, \cdots, n$$

and v_t satisfies

$$E(v_t) = 0, \; Var(v_t) = \sigma_v^2, \; Cov(v_t, v_s) = 0 \; \text{ for } t \neq s$$

Solutions for Autocorrelation: GLS [cont'd]

- **Idea:** let's transform the regression model so that the model has no autocorrelation.
  - Consider the previous period of the original model

$$Y_{t-1} = \beta_1 + \beta_2 X_{t-1} + e_{t-1}$$

  - Then pre-multiply the AR(1) coefficient ρ as

$$\rho Y_{t-1} = \rho\beta_1 + \rho\beta_2 X_{t-1} + \rho e_{t-1} \tag{2}$$

Solutions for Autocorrelation: GLS [cont'd]

- Now, subtract (2) from (1), it follows that

$$Y_t - \rho Y_{t-1} = \beta_1 (1-\rho) + \beta_2 (X_t - \rho X_{t-1}) + e_t - \rho e_{t-1}$$

- Define

$$\begin{aligned} Y_t^* &= Y_t - \rho Y_{t-1} \quad \text{for } t = 2, \cdots, n \\ X_{1t}^* &= 1 - \rho \quad \text{for } t = 2, \cdots, n \\ X_{2t}^* &= X_t - \rho X_{t-1} \quad \text{for } t = 2, \cdots, n \end{aligned}$$

- And we know that

$$v_t = e_t - \rho e_{t-1} \quad \text{for } t = 2, \cdots, n$$

then define $e_t^* = v_t$

Solutions for Autocorrelation: GLS [cont'd]

$$\implies \quad Y_t^* = \beta_1 X_{1t}^* + \beta_2 X_{2t}^* + e_t^* \quad \text{for } t = 2, \cdots, n \tag{3}$$

- Now, we know that the error terms in (3) are uncorrelated and satisfies the classical assumptions.

- **Transformation of the first observation** (*Prais-Winsten procedure*)
  - In fact, the Cochrane-Orcutt transformation holds for $t = 2, \cdots, n$ but not applicable to the first observation.
  - The Prais-Winsten procedure makes a reasonable transformation for $t = 1$.
  - The first observation in the regression model is

$$Y_1 = \beta_1 + \beta_2 X_1 + e_1$$

with the variance of e_1 is

$$Var(e_1) = \frac{\sigma_v^2}{1 - \rho}$$

- We want to make the variance of transformed regression error $\left(e_1^*\right)$ to be σ_v^2 which is $Var\left(e_t^*\right)$ in (3). That is, we multiply $\sqrt{1-\rho^2}$ as

$$\sqrt{1-\rho^2}Y_1 = \beta_1\sqrt{1-\rho^2} + \beta_2\sqrt{1-\rho^2}X_1 + \sqrt{1-\rho^2}e_1$$

$\Longrightarrow$

$$Y_1^* = \beta_1 X_{11}^* + \beta_2 X_{21}^* + e_1^* \tag{4}$$

where

$$Y_1^* = \sqrt{1-\rho^2}Y_1,\ X_{11}^* = \sqrt{1-\rho^2}$$
$$X_{21}^* = \sqrt{1-\rho^2}X_1,\ e_1^* = \sqrt{1-\rho^2}e_1$$

- **GLS estimator**: Consider the transformed model

$$Y_t^* = \beta_1 X_{1t}^* + \beta_2 X_{2t}^* + e_t^* \quad \text{for } t = 1, \cdots, n \tag{5}$$

given by (3) and (4). Then, OLS estimator in the transformed regression (5), which is the GLS estimator in the original model, is the **BLUE** (best linear unbiased estimator).

- **However!** We must know the value of ρ in the model which is a unknown parameter. $\rightarrow$ ***Infeasible GLS***

Solutions for Autocorrelation: FGLS

- **Feasible Generalized Least Squares (FGLS)**
  - We have to estimate ρ for the GLS transformation.
  - Consider AR(1) model as $e_t = \rho e_{t-1} + v_t$. Intuitive estimation of ρ will be OLS estimator of the given regression.
  - However, the regression errors e_t are unobservable, so we use the residual $\hat{e}_t$ instead of e_t.
  - Then the OLS estimates for ρ is given by

$$\hat{\rho} = \frac{\sum \hat{e}_t \hat{e}_{t-1}}{\sum \hat{e}_t^2}$$

- Once we get $\hat{\rho}$, then we can transform the model based on the transformation given by (3) and (4).
- **Cochrane-Orcutt estimation**: In order to reduce the uncertainty of $\hat{\beta}_1^*$ and $\hat{\beta}_2^*$, we repeat (again and again) the above procedure.
  - Draw new residuals based on $\hat{\beta}_1^*$ and $\hat{\beta}_2^*$.
  - Estimate new $\hat{\rho}$.
  - Transform the model using new $\hat{\rho}$.
  - $\cdots\cdots$