Remaining Topics in Time Series Analysis: VAR, VECM, Unit Root

Class 10 (Last Class)

Wonmun Shin

(wonmun.shin@sejong.ac.kr)

Department of Economics, Sejong University

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

VAR Model

Review on SEM

- SEM (Simultaneous Equation Model) appeared based on the insight that the single equation is not sufficient in understanding complicated economy.
  - **Structural-form** model: based on some underlying theoretical economic model
  - **Reduced-form** model: expressed as functions of exogenous variables
- In order to analyze impulse-responses of structural shock, we need to estimate the structural coefficients but there is endogeneity issue.
  - Therefore, we have learned ILS and 2SLS methodologies to earn consistent estimator for structural coefficients.
  - The key idea is that we should estimate reduced-form coefficients at first!

Basic Structure of VAR

- **VAR (Vector Auto-Regressive) model** is the AR model composed of vectors.
- From now, Y_t is $(n \times 1)$ vector:

$$Y_t = \begin{pmatrix} y_{1t} & y_{2t} & \cdots & y_{nt} \end{pmatrix}'$$

  - where $y_{1t}, y_{2t}, \cdots$ are stationary stochastic processes.

- General VAR(p) is given by:

$$\text{VAR}(p): \quad Y_t = \Phi_1 Y_{t-1} + \Phi_2 Y_{t-2} + \cdots + \Phi_{t-p} Y_{t-p} + e_t$$

- We will learn the simplest VAR, so let's suppose $n = 2$ and $p = 1$.

$$\text{VAR}(1): \quad Y_t = \Phi_1 Y_{t-1} + e_t$$

Structural Form

- Consider the following model:

$$y_{1t} = \beta_{12} y_{2t} + \gamma_{11} y_{1,t-1} + \gamma_{12} y_{2,t-1} + e_{1t}, \;\; e_{1t} \sim iidN\left(0, \sigma_1^2\right)$$
$$y_{2t} = \beta_{21} y_{1t} + \gamma_{21} y_{1,t-1} + \gamma_{22} y_{2,t-1} + e_{2t}, \;\; e_{2t} \sim iidN\left(0, \sigma_2^2\right)$$

  - y_{1t} and y_{2t} are endogenous variables $\rightarrow$ Endogeneity problem
  - We want to know the effect of structural shocks $(e_{1t},\, e_{2t})$ on endogenous variables $(y_{1t},\, y_{2t})$.
  - Note that the above equations are under-identified, then how do we know impulse-responses?

$$y_{1t} = \beta_{12} y_{2t} + \gamma_{11} y_{1,t-1} + \gamma_{12} y_{2,t-1} + e_{1t}$$
$$y_{2t} = \beta_{21} y_{1t} + \gamma_{21} y_{1,t-1} + \gamma_{22} y_{2,t-1} + e_{2t}$$

$$\Longrightarrow \quad y_{1t} - \beta_{12} y_{2t} = \gamma_{11} y_{1t-1} + \gamma_{12} y_{2t-1} + e_{1t}$$
$$-\beta_{21} y_{1t} + y_{2t} = \gamma_{21} y_{1t-1} + \gamma_{22} y_{2t-1} + e_{2t}$$

$$\Longrightarrow \begin{pmatrix} 1 & -\beta_{12} \\ -\beta_{21} & 1 \end{pmatrix} \begin{pmatrix} y_{1t} \\ y_{2t} \end{pmatrix} = \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \begin{pmatrix} y_{1t-1} \\ y_{2t-1} \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}$$

$$\Longrightarrow \quad B \cdot Y_t = \Gamma \cdot Y_{t-1} + e_t$$

Reduced Form

$$B \cdot Y_t = \Gamma \cdot Y_{t-1} + e_t$$

$$\implies B^{-1}BY_t = B^{-1}\Gamma Y_{t-1} + B^{-1}e_t$$

$$\implies Y_t = \Phi Y_{t-1} + u_t \text{ where } \Phi = B^{-1}\Gamma,\ u_t = B^{-1}e_t$$

- We can estimate Φ matrix because there is no endogeneity issue.
- Our final goal is estimating B matrix. Then how?
  - The residuals from the reduced-from estimation, $\hat{u}_t$.
  - And we know $u_t = B^{-1}e_t$, so we can estimate B^{-1} by imposing proper restrictions such as *Cholesky decomposition* of the covariance matrix of u_t.

Cointegration and VECM

Cointegration

- Recall that when a time series has to be differenced d times to make it stationary, we say that the time series is **integrated of order** d, $i.e.$ $Y_t \sim I(d)$.
- Generally, when $Y_{1t} \sim I(d)$ and $Y_{2t} \sim I(d)$, then $aY_{1t} + bY_{2t} \sim I(d)$ for constant a and b.
- But in some special cases, $aY_{1t} + bY_{2t} \sim I(d^*)$ where $d^* < d$! $\rightarrow$ We call the case **cointegration**.
- For example, suppose $X_t \sim I(1)$, and Y_{1t}, Y_{2t} are defined by:

$$Y_{1t} = 0.8X_t + e_t$$
$$Y_{2t} = X_t + u_t$$

  - Note that $Y_{1t} \sim I(1)$ and $Y_{2t} \sim I(1)$.
  - But if we consider $Y_{1t} - 0.8Y_{2t}$, then this linear combination is $I(0)$ process because the non-stationary component is removed by the linear combination.

VECM (Vector Error Correction Model)

- Consider we have two non-stationary time series (DSP) integrated of order 1.
  - Then we should take a difference on each series to make stationary VAR process.

$$\triangle Y_{1t} = \phi_{11} \triangle Y_{1t-1} + \phi_{12} \triangle Y_{2t-1} + e_{1t}$$
$$\triangle Y_{2t} = \phi_{21} \triangle Y_{2t-1} + \phi_{22} \triangle Y_{2t-1} + e_{2t}$$

  - The above one is the reduced-form VAR(1) with stationary processes.
- Interpretation of cointegration: **long-run relationship** of equilibrium between time-series data.
  - Imagine that Y_{1t} and Y_{2t} are cointegrated, for instance, $Y_{1t} - Y_{2t} \sim I(0)$.
  - Even if there is a long-run relationship between Y_{1t} and Y_{2t}, we did differencing without considering that relationship to make VAR(1).
  - In case, we should include the information about long-run relationship to equilibrium to correct error in the VAR model.

VECM (Vector Error Correction Model) [cont'd]

- **VECM (Vector Error Correction Model)**: Error-correction VAR in the process of cointegration.

$$\triangle Y_{1t} = \phi_{11} \triangle Y_{1t-1} + \phi_{12} \triangle Y_{2t-1} + \gamma_1 \left(Y_{1t} - Y_{2t} \right) + e_{1t}$$
$$\triangle Y_{2t} = \phi_{21} \triangle Y_{2t-1} + \phi_{22} \triangle Y_{2t-1} + \gamma_2 \left(Y_{1t} - Y_{2t} \right) + e_{2t}$$

  - $\gamma_1 \left(Y_{1t} - Y_{2t} \right)$, $\gamma_2 \left(Y_{1t} - Y_{2t} \right)$: ***Error-correction term***
  - γ_1 and γ_2 can be different in sign as well as size.
  - γ_1 and γ_2 represents the degree of ***speed of convergence*** to long-run equilibrium.

- How to know the existence of cointegration?
  - Based on intuition (or priori): long-run interest rate and short-run interest rate, stock prices and dividends
  - Cointegration test

Unit Root Test

Recall: Stationary Condition of ARMA model

- We have learned that stationary condition of ARMA model is related to the eigenvalues of the coefficient matrix of the state-space form.

$$|\lambda_1| < 1, |\lambda_2| < 1, |\lambda_3| < 1, \cdots, |\lambda_p| < 1.$$

- If there is at least one $\lambda_j \geq 1$, the process is non-stationary.
- Specifically, we are interested in $\lambda_j = 1$ which is a **"unit root"** of the characteristic equation.
- In other words, we say that a time-series data is non-stationary if it has a unit root.

Unit Root Test

- Consider a simple AR(1) model:

$$Y_t = \phi Y_{t-1} + e_t \quad e_t \sim iidN\left(0, \sigma^2\right)$$

- We will test the following:

$$\begin{cases} H_0: & \phi = 1 \text{ (unit root)} \\ H_1: & \phi < 1 \text{ (stationarity)} \end{cases}$$

- In order to test the above hypothesis, we should know the distribution of test statistic.

$$\frac{\hat{\phi} - 1}{\sqrt{Var\left(\hat{\phi}\right)}} \sim ??$$

Dickey-Fuller Test and Spurious Regression

- It is known that the t-statistic of the unit root test follows Dickey-Fuller distribution.
- Using the Dickey-Fuller distribution table, we can test the existence of unit root.
- If H_0 of unit root is not rejected, we decide that the process is non-stationary.

- What happens unless we detect unit root? $\rightarrow$ **Spurious regression**
  - For instance, there are two independent non-stationary process X_t, Y_t.
  - Since they are independent, if we regress Y_t on X_t ($Y_t = \beta X_t + e_t$), then the $\hat{\beta}$ should be insignificant.
  - However, if they are non-stationary it is possible that we have significant $\hat{\beta}$!

Augmented Dickey-Fuller Test

- In practice, we usually use **ADF (Augmented Dicky-Fuller) test** to detect unit roots.
- Let's start with AR(2) example:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t$$

- Here, the criterion for existence of unit root is $\phi_1 + \phi_2 = 1$.
  - Why? The characteristic equation for eigenvalues is $\lambda^2 - \phi_1\lambda - \phi_2 = 0$. Here, according to the properties of a quadratic equation, $\phi_1 = \lambda_1 + \lambda_2$, and $-\phi_2 = \lambda_1\lambda_2$. So, $\phi_1 + \phi_2 = \lambda_1 + \lambda_2 - \lambda_1\lambda_2$. Note that, if λ_1 or λ_2 is one, then $\phi_1 + \phi_2 = 1$.
- So we want to test:

$$\begin{cases} H_0: & \phi_1 + \phi_2 = 1 \\ H_1: & \phi_1 + \phi_2 < 1 \end{cases}$$

Augmented Dickey-Fuller Test [cont'd]

- Therefore, we need to modify the AR(2) model:

$$\begin{aligned} Y_t &= \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + e_t \\ &= \phi_1 Y_{t-1} + \phi_2 Y_{t-1} - \phi_2 Y_{t-1} + \phi_2 Y_{t-2} + e_t \\ &= (\phi_1 + \phi_2)\, Y_{t-1} - \phi_2 \triangle Y_{t-1} + e_t \\ &= \rho Y_{t-1} + \beta_1 \triangle Y_{t-1} + e_t \end{aligned}$$

  - by defining $\rho = \phi_1 + \phi_2$ and $\beta_1 = -\phi_2$

- Then, what we should test is:

$$\begin{cases} H_0: & \rho = 1 \\ H_1: & \rho < 1 \end{cases}$$

- And we can test the hypothesis using Dicky-Fuller distribution.

Augmented Dickey-Fuller Test [cont'd]

- In general, we can extend for AR(p):

$$Y_t = \rho Y_{t-1} + \beta_1 \triangle Y_{t-1} + \beta_2 \triangle Y_{t-2} + \cdots + \beta_{p-1} \triangle Y_{t-p+1} + e_t$$

  - where $\rho = \phi_1 + \phi_2 + \cdots + \phi_p$

- In practice, the equation for ADF test is:

$$\triangle Y_t = (\rho - 1) \, Y_{t-1} + \beta_1 \triangle Y_{t-1} + \beta_2 \triangle Y_{t-2} + \cdots + \beta_{p-1} \triangle Y_{t-p+1} + e_t$$

- However, we do not know how many lags should be considered for the unit root test (because we do not know the true p).
  - Therefore, the ADF test consider sufficient number of lags at the beginning and remove one-by-one from back (by testing the significance of $\hat{\beta}$).
  - So, the result of ADF test presents the p-value for $H_0 : \rho = 1$ as well as the optimal lag.