2 Vector Autoregressive (VAR) Model – Causality, Impulse Response Function and etc

Vector Autoregressive Process:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t,$$

where

$$y_t : k \times 1, \qquad \mu : k \times 1, \qquad \epsilon_t : k \times 1, \qquad \phi_i : k \times k.$$

Rewriting the above equation,

$$\phi(L)y_t = \mu + \epsilon_t,$$

where $\phi(L) = I_k - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$.

VAR(1) Model:

$$y_t = \phi_1 y_{t-1} + \epsilon_t, \qquad \text{i.e.,} \qquad (I_k - \phi_1 L)y_t = \epsilon_t.$$

When y_t is stationary, we obtain:

$$\begin{aligned}
y_t &= (I_k - \phi_1 L)^{-1}\epsilon_t \\
&= (I_k + \phi_1 L + \phi_1^2 L^2 + \phi_1^3 L^3 + \cdots)\epsilon_t \\
&= \epsilon_t + \phi_1\epsilon_{t-1} + \phi_1^2\epsilon_{t-2} + \phi_1^3\epsilon_{t-3} + \cdots
\end{aligned}$$

VAR(1)=VMA(∞)

VAR(2) Model:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \epsilon_t, \qquad \text{i.e.,} \qquad (I_k - \phi_1 L - \phi_2 L^2) y_{t-1} = \epsilon_t.$$

When y_t is stationary, we obtain:

$$y_{t-1} = (I_k - \phi_1 L - \phi_2 L^2)^{-1} \epsilon_t$$
$$= \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots$$

VAR(2)=VMA(∞)

95

VAR(p) Model:

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t,$$

i.e.,

$$(I_k - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p) y_{t-1} = \epsilon_t.$$

When y_t is stationary, we obtain:

$$
\begin{aligned}
y_t &= (I_k - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p)^{-1} \epsilon_t \\
&= \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots
\end{aligned}
$$

VAR(p)=VMA(∞)

2.1 Autocovariance Matrix and Autocorrelation Matrix

Let y_t be a $k \times 1$ vector.

Autocovariance Function Matrix:

$$\Gamma(\tau) = E((y_t - \mu)(y_{t-\tau} - \mu)'), \qquad \tau = 0, 1, 2, \cdots,$$

where $E(y_t) = \mu$. $\Gamma(\tau)$ is a $k \times k$ matrix.

$$\Gamma(\tau) = \Gamma(-\tau)'$$

Autocorrelation Function Matrix:

$$\rho(\tau) = D^{-1/2}\Gamma(\tau)D^{-1/2},$$

where the (i, j)th element of D is given by $\gamma_{ii}(\tau) = V(y_{it})$ for $i = j$ and zero otherwise.

$$\rho(\tau) = \rho(-\tau)'$$

2.2 Granger Cuasality Test (グレンジャー因果性テスト)

Consider a bivariate case.

Unrestricted Model (Sum of Squared Residuals, denoted by SSR_1):

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \phi_{11,1} & \phi_{12,1} \\ \phi_{21,1} & \phi_{22,1} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \cdots + \begin{pmatrix} \phi_{11,p} & \phi_{12,p} \\ \phi_{21,p} & \phi_{22,p} \end{pmatrix} \begin{pmatrix} y_{1,t-p} \\ y_{2,t-p} \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix}$$

$H_0: \phi_{12,1} = \phi_{12,2} = \cdots = \phi_{12,p} = 0$

When H_0 is correct, we say there is no causality from y_2 to y_1.

$\implies$ Granger Causality Test.

Restricted Model (Sum of Squared Residuals, denoted by SSR_0):

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \phi_{11,1} & 0 \\ \phi_{21,1} & \phi_{22,1} \end{pmatrix} \begin{pmatrix} y_{1,t-1} \\ y_{2,t-1} \end{pmatrix} + \cdots + \begin{pmatrix} \phi_{11,p} & 0 \\ \phi_{21,p} & \phi_{22,p} \end{pmatrix} \begin{pmatrix} y_{1,t-p} \\ y_{2,t-p} \end{pmatrix} + \begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix}$$

Asymptotically, we have the following distribution:

$$F = \frac{(SSR_0 - SSR_1)/p}{SSR_1/(T - 2p - 1)} \sim F(p, T - 2p - 1),$$

or

$$pF \sim \chi^2(p).$$

In general, we consider testing the Granger causality from y_j to y_i.

$$y_t = \mu + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t.$$

$$y_t : k \times 1, \qquad \mu : k \times 1, \qquad \phi_p : k \times k, \qquad \epsilon_t : k \times 1.$$

The null hypothesis is: $H_0 : \phi_{ij,1} = \phi_{ij,2} = \cdots = \phi_{ij,p} = 0.$
The alternative hypothesis is: $H_1 :$ not H_0.

SSR_0 = Sum of Squared Residuals under H_0
SSR_1 = Sum of Squared Residuals under H_1

Under H_0, the asymmptotic distribution is given by:

$$F = \frac{(\text{SSR}_0 - \text{SSR}_1)/p}{\text{SSR}_1/(T - kp - 1)} \sim F(p, T - kp - 1),$$

or

$$pF \sim \chi^2(p).$$

2.3 Impulse Response Function (インパルス応答関数):

$$\frac{\partial y_{i,t+k}}{\partial \epsilon_{j,t}}, \qquad k = 1, 2, \cdots,$$

where $i, j = 1, 2, \cdots, k$.

Example: AR(p) Process:

When y_t is stationary, we obtain:

$$y_t = (I_k - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p)^{-1} \epsilon_t$$
$$= \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots$$

$$\frac{\partial y_{i,t+k}}{\partial \epsilon_{j,t}} = \theta_{ij,k}, \qquad k = 1, 2, \cdots,$$

where $\theta_{ij,k}$ denotes the (i, j)th element of θ_k.

3 Unit Root (単位根) and Cointegration (共和分)

3.1 Unit Root (単位根)

1. Why is a unit root problem important?

   (a) Economic variables increase over time in general.

   One of the assumptions of OLS is stationarity on y_t and x_t.

   This assumption implies that $\frac{1}{T}X'X$ converges to a fixed matrix as T is large.

That is, asymptotic normality of OLS estimator goes not hold.

(b) In nonstationary time series, the unit root is the m ost important.

In the case of unit root, OLSE of the first-order autoregressive coefficient is consistent.

OLSE is $\sqrt{T}$-consistent in the case of stationary AR(1) process, but OLSE is T-consistent in the case of nonstationay AR(1) process.

(c) A lot of economic variables increase over time.

It is important to check an economic variable is trend stationary (i.e., $y_t = a_0 + a_1 t + \epsilon_t$) or difference stationary (i.e., $y_t = b_0 + y_{t-1} + \epsilon_t$).

Consider *k*-step ahead prediction for both cases.

$$\text{(Trend Stationarity)} \qquad y_{t+k|t} = a_0 + a_1(t + k)$$

$$\text{(Difference Stationarity)} \qquad y_{t+k|t} = b_0 k + y_t$$

2. The Case of $|\phi_1| < 1$:

$$y_t = \phi_1 y_{t-1} + \epsilon_t, \qquad \epsilon_t \sim \text{i.i.d. } N(0, \sigma_\epsilon^2), \qquad y_0 = 0, \qquad t = 1, \cdots, T$$

Then, OLSE of ϕ_1 is:

$$\hat{\phi}_1 = \frac{\displaystyle\sum_{t=1}^{T} y_{t-1} y_t}{\displaystyle\sum_{t=1}^{T} y_{t-1}^2}.$$

In the case of $|\phi_1| < 1$,

$$\hat{\phi}_1 = \phi_1 + \frac{\dfrac{1}{T}\displaystyle\sum_{t=1}^{T} y_{t-1}\epsilon_t}{\dfrac{1}{T}\displaystyle\sum_{t=1}^{T} y_{t-1}^2} \longrightarrow \phi_1 + \frac{E(y_{t-1}\epsilon_t)}{E(y_{t-1}^2)} = \phi_1.$$

Note as follows:
$$\frac{1}{T} \sum_{t=1}^{T} y_{t-1}\epsilon_t \longrightarrow \mathrm{E}(y_{t-1}\epsilon_t) = 0.$$

By the central limit theorem,
$$\frac{\overline{y\epsilon} - \mathrm{E}(\overline{y\epsilon})}{\sqrt{\mathrm{V}(\overline{y\epsilon})}} \longrightarrow N(0, 1)$$

where
$$\overline{y\epsilon} = \frac{1}{T} \sum_{t=1}^{T} y_{t-1}\epsilon_t.$$

$$
\begin{aligned}
\mathrm{E}(\overline{y\epsilon}) &= 0, \\
\mathrm{V}(\overline{y\epsilon}) &= \mathrm{V}\Big(\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\epsilon_t\Big) = \mathrm{E}\Big(\big(\frac{1}{T}\sum_{t=1}^{T} y_{t-1}\epsilon_t\big)^2\Big) \\
&= \frac{1}{T^2}\mathrm{E}\Big(\sum_{t=1}^{T}\sum_{s=1}^{T} y_{t-1} y_{s-1}\epsilon_t\epsilon_s\Big) = \frac{1}{T^2}\mathrm{E}\Big(\sum_{t=1}^{T} y_{t-1}^2 \epsilon_t^2\Big) = \frac{1}{T}\sigma^2\gamma(0).
\end{aligned}
$$

Therefore,

$$
\frac{\overline{y\epsilon}}{\sqrt{\sigma^2\gamma(0)/T}} = \frac{1}{\sigma_\epsilon\sqrt{\gamma(0)}}\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{t-1}\epsilon_t \longrightarrow N(0,1),
$$

which is rewritten as:

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} y_{t-1} \epsilon_t \longrightarrow N(0, \sigma_\epsilon^2 \gamma(0)).$$

Using $\dfrac{1}{T} \sum_{t=1}^{T} y_{t-1}^2 \longrightarrow E(y_{t-1}^2) = \gamma(0)$, we have the following asymptotic distribution:

$$\sqrt{T}(\hat{\phi}_1 - \phi_1) = \frac{\dfrac{1}{\sqrt{T}} \sum_{t=1}^{T} y_{t-1} \epsilon_t}{\dfrac{1}{T} \sum_{t=1}^{T} y_{t-1}^2} \longrightarrow N\left(0, \frac{\sigma_\epsilon^2}{\gamma(0)}\right) = N\left(0, 1 - \phi_1^2\right).$$

Note that $\gamma(0) = \dfrac{\sigma_\epsilon^2}{1 - \phi_1^2}$.

3. In the case of $\phi_1 = 1$, as expected, we have:

$$\sqrt{T}(\hat{\phi}_1 - 1) \longrightarrow 0.$$

That is, $\hat{\phi}_1$ has the distribution which converges in probability to $\phi_1 = 1$ (i.e., degenerated distribution).

Is this true?

4. **The Case of $\phi_1 = 1$:** $\implies$ Random Walk Process

$y_t = y_{t-1} + \epsilon_t$ with $y_0 = 0$ is written as:

$$y_t = \epsilon_t + \epsilon_{t-1} + \epsilon_{t-2} + \cdots + \epsilon_1.$$

Therefore, we can obtain:

$$y_t \sim N(0, \sigma_\epsilon^2 t).$$

The variance of y_t depends on time t. $\implies y_t$ is nonstationary.

5. Remember that $\hat{\phi}_1 = \phi_1 + \dfrac{\sum y_{t-1}\epsilon_t}{\sum y_{t-1}^2}$.

   (a) First, consider the numerator $\sum y_{t-1}\epsilon_t$.

We have $y_t^2 = (y_{t-1} + \epsilon_t)^2 = y_{t-1}^2 + 2y_{t-1}\epsilon_t + \epsilon_t^2$.

Therefore, we obtain:

$$y_{t-1}\epsilon_t = \frac{1}{2}(y_t^2 - y_{t-1}^2 - \epsilon_t^2).$$

Taking into account $y_0 = 0$, we have:

$$\sum_{t=1}^{T} y_{t-1}\epsilon_t = \frac{1}{2}y_T^2 - \frac{1}{2}\sum_{t=1}^{T}\epsilon_t^2.$$

Divided by $\sigma_\epsilon^2 T$ on both sides, we have the following:

$$\frac{1}{\sigma_\epsilon^2 T}\sum_{t=1}^{T} y_{t-1}\epsilon_t = \frac{1}{2}\left(\frac{y_T}{\sigma_\epsilon \sqrt{T}}\right)^2 - \frac{1}{2\sigma_\epsilon^2}\frac{1}{T}\sum_{t=1}^{T}\epsilon_t^2.$$

From $y_t \sim N(0, \sigma_\epsilon^2 t)$, we obtain the following result:

$$\left(\frac{y_T}{\sigma_\epsilon \sqrt{T}}\right)^2 \sim \chi^2(1).$$

Moreover, the second term is derived from:

$$\frac{1}{T} \sum_{t=1}^{T} \epsilon_t^2 \longrightarrow \sigma_\epsilon^2.$$

Therefore,

$$\frac{1}{\sigma_\epsilon^2 T} \sum_{t=1}^{T} y_{t-1}\epsilon_t = \frac{1}{2}\left(\frac{y_T}{\sigma \sqrt{T}}\right)^2 - \frac{1}{2\sigma_\epsilon^2}\frac{1}{T} \sum_{t=1}^{T} \epsilon_t^2 \longrightarrow \frac{1}{2}(\chi^2(1) - 1).$$

(b) Next, consider $\sum y_{t-1}^2$.

$$\mathrm{E}\left(\sum_{t=1}^{T} y_{t-1}^2\right) = \sum_{t=1}^{T} \mathrm{E}(y_{t-1}^2) = \sum_{t=1}^{T} \sigma_\epsilon^2(t-1) = \sigma_\epsilon^2 \frac{T(T-1)}{2}.$$

Thus, we obtain the following result:

$$\frac{1}{T^2}\mathrm{E}\left(\sum_{t=1}^{T} y_{t-1}^2\right) \longrightarrow \text{ a fixed value.}$$

Therefore,

$$\frac{1}{T^2} \sum_{t=1}^{T} y_{t-1}^2 \longrightarrow \text{ a distribution.}$$

116