Econometrics 1, Class Week 10

Back-up Video: Class week 10 video (click here)

Learning Outcomes

(a) Vector Autoregressive (VAR) processes.

(b) Forecasting.

(c) Granger causality.

(d) Vector Error Correction Model (VECM) representation.

(e)-(f) Cointegration.

(g) ARDL model implied by VAR.

Prerequisites

1. Concepts and results of classes of weeks 7, 8 and 9.

2. Multivariate normal distribution (Distributional Handout): conditional vs. unconditional normal distribution.

(a) VAR(1) Setting

Consider the vector-valued process $\{\mathbf{y}_t\} = \{(y_{1t}, y_{2t})'\}$ which is assumed to be autoregressive of order one (VAR(1)),

$$\mathbf{y}_t = \mathbf{A}_0 + \mathbf{A}_1 \mathbf{y}_{t-1} + \epsilon_t \;\Leftrightarrow\; (\mathbf{I}_2 - \mathbf{A}_1 L)\mathbf{y}_t = \mathbf{A}_0 + \epsilon_t,$$

for any t, where

$$\mathbf{A}_0 = \begin{bmatrix} a_1^0 \\ a_2^0 \end{bmatrix}, \; \mathbf{A}_1 = \begin{bmatrix} a_{11}^1 & a_{12}^1 \\ a_{21}^1 & a_{22}^1 \end{bmatrix},$$

and ϵ_t is multivariate WN, i.e.

$$\epsilon_t = (\epsilon_{1t}, \epsilon_{2t})' \overset{i.i.d.}{\sim} N(\mathbf{0}, \Omega), \text{ with } \Omega = \begin{bmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{bmatrix} \text{ p.d.s.}$$

Written out in single equation form,

$$\begin{aligned} y_{1t} &= a_1^0 + a_{11}^1 y_{1,t-1} + a_{12}^1 y_{2,t-1} + \epsilon_{1t}, \\ y_{2t} &= a_2^0 + a_{21}^1 y_{1,t-1} + a_{22}^1 y_{2,t-1} + \epsilon_{2t}. \end{aligned}$$

Suppose data $\{\mathbf{y}_t, t = 1, \cdots, T\}$ are available for periods 1 to T. On the basis of such data, the parameters can be estimated by OLS:

(i) $a_1^0, a_{11}^1, a_{12}^1$ from an OLS regression of y_{1t} on a constant, $y_{1,t-1}$ and $y_{2,t-1}$;

(ii) $a_2^0, a_{21}^1, a_{22}^1$ from an OLS regression of y_{2t} on a constant, $y_{1,t-1}$ and $y_{2,t-1}$;

(iii) and ω_{ij} from the OLS residuals $\{\hat{\epsilon}_{it}, t = 1, 2, t = 1, \cdots, T\}$, as $\hat{\omega}_{ij} = \frac{1}{T}\sum_{t=1}^{T} \hat{\epsilon}_{it}\hat{\epsilon}_{jt}$, $i, j \in \{1, 2\}$.

(b) Forecasting

Analogous to process (1) in the assignment for week 7, the first-order Markov property of the process $\{\mathbf{y}_t\}$ implies

$$\begin{aligned}
\hat{\mathbf{y}}_{T+1|T} &= \mathbb{E}[\mathbf{y}_{T+1}|\mathbf{y}_{t\leq T}] \\
&= \mathbb{E}[\mathbf{y}_{T+1}|\mathbf{y}_{T}] = \mathbf{A}_0 + \mathbf{A}_1\mathbf{y}_T, \\
\hat{\mathbf{y}}_{T+2|T} &= \mathbb{E}[\mathbf{y}_{T+2}|\mathbf{y}_{t\leq T}] \\
&= \mathbf{A}_0 + \mathbf{A}_1\mathbb{E}[\mathbf{y}_{T+1}|\mathbf{y}_{T}] \\
&= (\mathbf{I}_2 + \mathbf{A}_1)\mathbf{A}_0 + \mathbf{A}_1^2\mathbf{y}_T,
\end{aligned}$$

where $\mathbf{I}_2$ is the identity matrix of dimension 2. These conditional expectations can be estimated by replacing $\mathbf{A}_0$ and $\mathbf{A}_1$ by their estimates from part (a).

(c) Granger Causality

Granger causality relates to usefulness in prediction.

y_{1t} is not Granger-causal with respect to y_{2t} if lagged values of y_{1t} do not help in the prediction of y_{2t}, i.e. when $a_{21}^{1} = 0$.

Typically, economic causality and Granger causality run in opposite directions.

Example: A listed firm's earning determines the dividends its stockholder receive; and its dividends Granger cause the firm's earnings.

(d) VECM

Analogous to ECM studied in week 9.

The form in (d) is a vector error correction model (VECM) which follows from

$$
\begin{aligned}
\Delta \mathbf{y}_t &= \mathbf{A}_0 + (\mathbf{A}_1 - \mathbf{I}_2)\mathbf{y}_{t-1} + \epsilon_t \\
\Rightarrow \Pi &= \mathbf{A}_1 - \mathbf{I}_2 \\
\Delta \mathbf{y}_t &= \mathbf{A}_0 + \Pi \mathbf{y}_{t-1} + \epsilon_t.
\end{aligned}
$$

(e) Cointegration

General principle: In VECM, order of integration of the LHS = order of integration of the RHS. And note: In this example, the rank of Π can be 0, 1 or 2. Consider three cases:

(iii) y_{it}, $i = 1, 2$, are both $I(1)$, i.e. contain a unit root, and not cointegrated. Then, the term $\Pi \mathbf{y}_t$ in the VECM must disappear for all realizations of $\mathbf{y}_t$. Hence, $\Pi = \mathbf{0}$, i.e. $\mathbf{A}_1 = \mathbf{I}_2$, and so $\mathrm{rk}(\Pi) = 0$.

(ii) y_{it}, $i = 1, 2$, are both $I(1)$, i.e. contain a unit root, and cointegrated. Then, rk(Π) equals the number of cointegrating vectors, so in this case rk(Π) $= 1$.

To see that unit roots imply rank deficiency of Π, note that

$$\begin{aligned} 0 &= |\mathbf{I}_2 - \mathbf{A}_1 z| \\ &= \left| \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} a_{11}^1 z & a_{12}^1 z \\ a_{21}^1 z & a_{22}^1 z \end{bmatrix} \right| \\ &= \begin{vmatrix} 1 - a_{11}^1 z & -a_{12}^1 z \\ -a_{21}^1 z & 1 - a_{22}^1 z \end{vmatrix} \\ &= (1 - a_{11}^1 z)(1 - a_{22}^1 z) - a_{12}^1 a_{21}^1 z^2, \end{aligned}$$

so a unit root, i.e. $z = 1$ solving the preceding equality, implies

$$(1 - a_{11}^1)(1 - a_{22}^1) - a_{12}^1 a_{21}^1 = 0,$$

which is equivalent to $\frac{a_{12}^1}{1-a_{11}^1} = \frac{1-a_{22}^1}{a_{21}^1}$ $(\star)$. But this implies that

$$\text{rk}(\Pi) = \text{rk}\left(\begin{bmatrix} a_{11}^1 - 1 & a_{12}^1 \\ a_{21}^1 & a_{22}^1 - 1 \end{bmatrix} \right) = 1.$$

(Multiplying the first row by $\frac{a_{21}^1}{a_{11}^1 - 1}$ yields $[a_{21}^1, \frac{a_{21}^1 a_{12}^1}{a_{11}^1 - 1}]$, which is equal to the second row, using the equality $(\star)$.)

In this case, Π can be expressed as $\Pi = \beta\mathbf{r}'$, for $\beta, \mathbf{r} \in \mathbb{R}^2$, where $\mathbf{r}$ is a cointegrating vector, i.e. the scalar random variable $z_t = \mathbf{r}'\mathbf{y}_t$, a linear combination of y_{1t} and y_{2t}, is $I(0)$. Then,

$$\Delta\mathbf{y}_t = \mathbf{A}_0 + \beta z_{t-1} + \epsilon_t.$$

(i) y_{it}, $i = 1, 2$, are both stationary. Then, $|\mathbf{I}_2 - \mathbf{A}_1 z| = 0$ has both roots outside the unit circle, and Π has full rank 2.

Informal argument: To see this, suppose *to the contrary* that $\text{rk}(\Pi) < 2$:

$$\begin{aligned}
\exists \mathbf{w} \in \mathbb{R}^2 : \mathbf{w}'\Pi &= \mathbf{0}' \\
\Rightarrow \qquad \mathbf{w}'\Delta\mathbf{y}_t &= \mathbf{w}'\mathbf{A}_0 + \mathbf{w}'\Pi\mathbf{y}_{t-1} + \mathbf{w}'\epsilon_t = \mathbf{w}'\mathbf{A}_0 + \mathbf{w}'\epsilon_t \\
\Rightarrow \qquad \mathbf{w}'\Delta\mathbf{y}_t & \quad \text{is } I(0), \text{ i.e. stationary.}
\end{aligned}$$

The last implication is consistent with $\mathbf{w}'\mathbf{y}_t$ being $I(1)$, i.e. at least one series contains a unit root. This contradicts that y_{it}, $i = 1, 2$, are both stationary. So stationary requires that Π have full rank.

(f) VECM s.t. Cointegration Restriction

This is case (ii) in part (e) above, with $\Pi = \beta \mathbf{r}'$ for some $\beta, \mathbf{r} \in \mathbb{R}^2$. Note: For any non-zero scalar Q:

$$\Pi = \beta \mathbf{r}' = \beta Q Q^{-1} \mathbf{r}' = \tilde{\beta} \tilde{\mathbf{r}}'$$

for $\tilde{\beta} = \beta Q$ and $\tilde{\mathbf{r}}' = Q^{-1}\mathbf{r}'$, so that $\tilde{\mathbf{r}}$ is also a cointegrating vector. Hence, the cointegrating vector is not identified (unique), unless a restriction is imposed (which is equivalent of choosing a specific Q).

One possible restriction is to have $\mathbf{r}$ normalized to have its first component equal to one. Then, $\mathbf{r}' = [1, b]$ for some b, so that $z_t = \mathbf{r}'\mathbf{y}_t = y_{1t} + by_{2t}$. From above,

$$\begin{aligned}
\Delta y_{1t} &= a_1^0 + (a_{11}^1 - 1)y_{1,t-1} + a_{12}^1 y_{2,t-1} + \epsilon_{1t} \\
&= a_1^0 + (a_{11}^1 - 1)\left(y_{1,t-1} + \frac{a_{12}^1}{a_{11}^1 - 1} y_{2,t-1}\right) + \epsilon_{1t} \\
\Delta y_{2t} &= a_2^0 + a_{21}^1 y_{1,t-1} + (a_{22}^1 - 1)y_{2,t-1} + \epsilon_{2t} \\
&= a_2^0 + a_{21}^1 \left(y_{1,t-1} + \frac{a_{22}^1 - 1}{a_{21}^1} y_{2,t-1}\right) + \epsilon_{1t},
\end{aligned}$$

and from the unit root condition that implies the rank deficiency of Π, it follows that

$$b = \frac{a_{12}^1}{a_{11}^1 - 1} = \frac{a_{22}^1 - 1}{a_{21}^1},$$

while $\beta = [a_{11}^1 - 1, a_{21}^1]'$.

ARDL implied by VAR

Starting from

$$y_{1t} = a_1^0 + a_{11}^1 y_{1,t-1} + a_{12}^1 y_{2,t-1} + \epsilon_{1t},$$

with $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t})'$ satisfying $\epsilon_t \sim N(\mathbf{0}, \Omega)$, the conditional distribution of ϵ_{1t}, given ϵ_{2t}, is

$$\epsilon_{1t}|\epsilon_{2t} \sim N\left(\frac{\omega_{12}}{\omega_{22}}\epsilon_{2t}, \omega_{11} - \frac{\omega_{12}^2}{\omega_{22}}\right),$$

so that $\epsilon_{2t} = y_{2t} - a_2^0 - a_{21}^1 y_{1,t-1} - a_{22}^1 y_{2,t-1}$ implies

$$\begin{aligned}
\mathbb{E}[y_{1t}|y_{2t}, \mathbf{y}_{t-1}] &= a_1^0 + a_{11}^1 y_{1,t-1} + a_{12}^1 y_{2,t-1} + \mathbb{E}[\epsilon_{1t}|y_{2t}, \mathbf{y}_{t-1}] \\
&= a_1^0 + a_{11}^1 y_{1,t-1} + a_{12}^1 y_{2,t-1} \\
&\quad + \frac{\omega_{12}}{\omega_{22}}\left(y_{2t} - a_2^0 - a_{21}^1 y_{1,t-1} - a_{22}^1 y_{2,t-1}\right) \\
&= a_1^0 - \frac{\omega_{12}}{\omega_{22}}a_2^0 + \frac{\omega_{12}}{\omega_{22}}y_{2t} + \left(a_{11}^1 - \frac{\omega_{12}}{\omega_{22}}a_{21}^1\right)y_{1,t-1} \\
&\quad + \left(a_{12}^1 - \frac{\omega_{12}}{\omega_{22}}a_{22}^1\right)y_{2,t-1}.
\end{aligned}$$

Therefore, the ARDL(1,1) model $y_{1t} = \alpha_0 + \beta_0 y_{2t} + \beta_1 y_{2,t-1} + \alpha_1 y_{1,t-1} + u_t$ has

$$\begin{aligned}
\alpha_0 &= a_1^0 - \frac{\omega_{12}}{\omega_{22}}a_2^0,\ \alpha_1 = a_{11}^1 - \frac{\omega_{12}}{\omega_{22}}a_{21}^1, \\
\beta_0 &= \frac{\omega_{12}}{\omega_{22}},\ \beta_1 = a_{12}^1 - \frac{\omega_{12}}{\omega_{22}}a_{22}^1
\end{aligned}$$

and $u_t \overset{i.i.d.}{\sim} N\left(0, \omega_{11} - \omega_{12}^2/\omega_{22}\right)$.