CH4 Vector Autoregression (VAR) Models

Intervention Model
Transfer Function Model
VAR Model in Reduced Form: Stationary
Impulse Response Function
Forecast Error Variance Decomposition
Multi-equation VAR in Reduced Form
Granger Causality
Structural VAR

Intervention Model

$$\boxed{y_t = a_0 + a_1 y_{t-1} + c_0 z_t + \varepsilon_t, \ |a_1| < 1,}$$

a **deterministic** intervention $z_t = \begin{cases} 0, t < T_0 \\ 1, t \geq T_0 \end{cases}$

The pulse response function is

$$y_{T_0+t} = \frac{a_0}{1-a_1} + c_0 \sum_{i=0}^{\infty} a_1^i z_{T_0+t-i} + \sum_{i=0}^{\infty} a_1^i \varepsilon_{T_0+t-i}.$$

The long-run effect of the intervention ($t \to \infty$) is

$$\frac{c_0}{1-a_1} = \frac{c_0 + a_0}{1-a_1} - \frac{a_0}{1-a_1}$$

Intervention Model: Extension

$$\boxed{y_t = a_0 + A(L)y_{t-1} + c_0 z_t + B(L)\varepsilon_t \text{ (ARMA(p,q) intervention model)}}$$

$$\boxed{y_t = a_0 + A(L)y_{t-1} + c_0 z_{t-d} + B(L)\varepsilon_t \text{ (ARMA(p,q) delayed intervention model)}}$$

$$z_t = \begin{cases} 1, & t = T_0 \\ 0, & \text{otherwise} \end{cases}$$

(Pulse function)

$$z_t = \begin{cases} 0, & t \le T_0 - 1 \\ 1/4, & t = T_0, \\ 1/2, & t = T_0 + 1 \\ 3/4, & t = T_0 + 2 \\ 1, & t \ge T_0 + 3 \end{cases} \quad \text{(Gradually changing function)}$$

$$z_t = \begin{cases} 0, & t \le T_0 - 1 \\ 1, & t = T_0, \\ 3/4, & t = T_0 + 1 \\ 1/2, & t = T_0 + 2 \\ 1/4, & t = T_0 + 3 \\ 0, & t \ge T_0 + 4 \end{cases} \quad \text{(Prolonged impulse function)}$$

Intervention Model: Estimation

Estimation: First, estimate the most appropriate models for both the pre- and post-intervention periods to check if the coefficients in the model are invariant to the intervention. If no, estimate the various models over the entire sample period and perform diagnostic checks of the estimated model to ensure that: (1) All coefficients should be significant and the AR coefficients imply that the series is stationary; (2) The residuals should approximate white noise; (3) The selected model outperforms other alternatives: using the AIC, SBC. Three steps for the estimation: P244-246.

Note: The effects of the intervention will change if $\{y_t\}$ has a unit root. In this case, a pulse intervention will have a permanet effect on the level of $\{y_t\}$; a pure jump intervention will act as a drift term in the process. An intervention will have a temporary effect on a unit root process if all values of $\{z_t\}$ sum to zero.

Transfer Function Model

$$\boxed{y_t = a_0 + A(L)y_{t-1} + C(L)z_t + B(L)\varepsilon_t}$$ a stochastic exogenous variable z_t

$E(z_t\varepsilon_{t-s}) = 0$ for all s and t; $\{z_t\}$ are independent.

$C(L)$ is called the transfer function.

The cross correlation function (**CCF**) between y_t and the various z_{t-i} is

$$\boxed{\rho_{yz}(i) = \frac{Cov(y_t, z_{t-i})}{\sigma_y \sigma_z}.}$$

The cross-covariance function (**CCVF**) between y_t and the various z_{t-i} is

$$\boxed{\gamma_{yz}(i) = \frac{Cov(y_t, z_{t-i})}{\sigma_z^2}.}$$

Transfer Function Model: Example

$y_t = a_1 y_{t-1} + c_d z_{t-d} + \varepsilon_t$, where z_t is i.i.d. with $E z_t = 0$ and $Var(z_t) = \sigma_z^2$.

$$\begin{aligned} y_t &= c_d z_{t-d}/(1 - a_1 L) + \varepsilon_t/(1 - a_1 L) \\ &= c_d \sum_{i=0}^{\infty} a_1^i z_{t-d-i} + \varepsilon_t/(1 - a_1 L), \end{aligned}$$

$E y_t z_t = 0, E y_t z_{t-1} = 0, \cdots, E y_t z_{t-d+1} = 0$

$E y_t z_{t-d} = c_d \sigma_z^2$

$E y_t z_{t-d-1} = c_d a_1 \sigma_z^2$

$E y_t z_{t-d-2} = c_d a_1^2 \sigma_z^2$

$$\begin{aligned} \rho_{yz}(i) &= 0, \quad i \le d-1, \\ &= c_d a_1^{i-d} \sigma_z/\sigma_y, \quad i \ge d. \end{aligned}$$

$$\begin{aligned} \gamma_{yz}(i) &= 0, \quad i \le d-1, \\ &= c_d a_1^{i-d}, \quad i \ge d. \end{aligned}$$

$\rho_{yz}(i) = a_1 \rho_{yz}(i-1), \quad i \ge d+1$ (decay at the rate a_1)

Transfer Function Model

$y_t = a_1 y_{t-1} + c_d z_{t-d} + c_{d+1} z_{t-d-1} + \varepsilon_t$, where $Ez_t = 0$ and $Var(z_t) = \sigma_z^2$.

$$
\begin{aligned}
\rho_{yz}(i) &= 0, \quad i \leq d-1, \\
&= c_d \sigma_z / \sigma_y, \; i = d, \\
&= (c_d a_1 + c_{d+1}) \sigma_z / \sigma_y, \; i = d+1, \\
&= (c_d a_1 + c_{d+1}) a_1^{i-d-1} \sigma_z / \sigma_y, \; i \geq d+2.
\end{aligned}
$$

$$
\begin{aligned}
\gamma_{yz}(i) &= 0, \quad i \leq d-1, \\
&= c_d, \; i = d, \\
&= c_d a_1 + c_{d+1}, \; i = d+1, \\
&= (c_d a_1 + c_{d+1}) a_1^{i-d-1}, \; i \geq d+2.
\end{aligned}
$$

$\rho_{yz}(i) = a_1 \rho_{yz}(i-1), \;\; i \geq d+2$ (decay at the rate a_1)

Transfer Function Model

$$y_t = a_0 + A(L)y_{t-1} + C(L)z_t + B(L)\varepsilon_t.$$

$\gamma_{yz}(i) = 0$, until the first nonzero element of $C(L)$

$B(L)$ is immaterial to CCVF

A spike in the CCVF indicates a nonzero element of $C(L)$

All spikes have a decay pattern.

$y_t = a_1 y_{t-1} + a_2 y_{t-2} + c_d z_{t-d} + \varepsilon_t$. The CCVF satisfies

$$\begin{aligned}
\gamma_{yz}(i) &= 0,\ i \le d-1, \\
&= c_d,\ i = d, \\
&= a_1\gamma_{yz}(i-1) + a_2\gamma_{yz}(i-2),\ i \ge d+1.
\end{aligned}$$

———an initial spike at lag d, then the decay pattern.

Restriction: 1) Restrict the form of the transfer function. 2) No feedback from $\{y_t\}$ to $\{z_t\}$. For the coefficients of $C(L)$ to be unbiased estimates of the impact effects of $\{z_t\}$ on $\{y_t\}$, z_t must be uncorrelated with the error term $\{\varepsilon_t\}$ at all leads and lags.

Vector White Noise and Stationary

Vector White Noise Process

$\{\varepsilon_t\}$ satisfying: (i) $E\varepsilon_t = \mathbf{0}$; (ii) $E(\varepsilon_t\varepsilon_t') = \Omega$ is a definite matrix; (iii) $E(\varepsilon_t\varepsilon_s') = \mathbf{0}$, $t \neq s$.

Let $y_t = (y_{1t}, \cdots, y_{nt})'$. The mean of y_t is

$$Ey_t = (Ey_{1t}, \cdots, Ey_{nt})' = (\mu_{1t}, \cdots, \mu_{nt})' \equiv \mu_t'.$$

The autocovariance matrix (or function) is

$$\begin{aligned}\Gamma(t,0) &= E\left[(y_t-\mu_t)(y_t-\mu_t)'\right] \\ &= \begin{pmatrix} var(y_{1t}) & cov(y_{1t},y_{2t}) & \cdots & cov(y_{1t},y_{nt}) \\ cov(y_{2t},y_{1t}) & var(y_{2t}) & \cdots & cov(y_{2t},y_{nt}) \\ \vdots & \vdots & \vdots & \vdots \\ cov(y_{nt},y_{1t}) & cov(y_{nt},y_{1t}) & \cdots & var(y_{nt}) \end{pmatrix}\end{aligned}$$

$$\begin{aligned}\Gamma(t,h) &= E\left[(y_t-\mu_t)(y_{t-h}-\mu_{t-h})'\right] \\ &= \begin{pmatrix} cov(y_{1t},y_{1,t-h}) & cov(y_{1t},y_{2,t-h}) & \cdots & cov(y_{1t},y_{n,t-h}) \\ cov(y_{2t},y_{1,t-h}) & cov(y_{1t},y_{2,t-h}) & \cdots & cov(y_{2t},y_{n,t-h}) \\ \vdots & \vdots & \vdots & \vdots \\ cov(y_{nt},y_{1,t-h}) & cov(y_{nt},y_{2,t-h}) & \cdots & cov(y_{nt},y_{n,t-h}) \end{pmatrix}\end{aligned}$$

$\{y_t\}$ is **stationary** if the second moment is finite and $Ey_t = \mu$

and $\Gamma(t,h)$ is related with h, but not with t.

Vector White Noise and Stationary

Example: Let $\left\{ \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix} \right\}$ is a vector white process with var-covariance matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Show that $\left\{ \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix} \right\}$ is also a vector white process, where $\varepsilon_{1t} = e_{1t} + 2e_{2t}$, $\varepsilon_{2t} = e_{2t}$.

Vector Autoregression Model

Both $\{y_t\}$ and $\{z_t\}$ are endogenous.

y_t and z_t are allowed to affect each other (feedback effect).

Consider the simple bivariate system

$$\boxed{\textit{the structural VAR}: \begin{cases} y_t = b_{10} - b_{12}z_t + \gamma_{11}y_{t-1} + \gamma_{12}z_{t-1} + \varepsilon_{yt} \\ z_t = b_{20} - b_{21}y_t + \gamma_{21}y_{t-1} + \gamma_{22}z_{t-1} + \varepsilon_{zt} \end{cases}}$$

where $\{y_t\}$ and $\{z_t\}$ are stationary, ε_{yt} and ε_{zt} are white-noise with variances σ_y^2 and σ_z^2, respectively, and ε_{yt} and ε_{zt} are uncorrelated. Write the structural model as

$$\begin{pmatrix} 1 & b_{12} \\ b_{21} & 1 \end{pmatrix} \begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} b_{10} \\ b_{20} \end{pmatrix} + \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} \begin{pmatrix} y_{t-1} \\ z_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{pmatrix},$$

denoted as

$$x_t = A_0 + A_1 x_{t-1} + e_t,$$

$$e_t = \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix} = \begin{pmatrix} 1 & b_{12} \\ b_{21} & 1 \end{pmatrix}^{-1} \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{pmatrix} = \begin{pmatrix} (\varepsilon_{yt} - b_{12}\varepsilon_{zt}) / (1 - b_{12}b_{21}) \\ (\varepsilon_{zt} - b_{21}\varepsilon_{yt}) / (1 - b_{12}b_{21}) \end{pmatrix}.$$

Vector Autoregression Model

we obtain the VAR in the reduced form:

$$\text{VAR in standard form: } \begin{cases} y_t = a_{10} + a_{11}y_{t-1} + a_{12}z_{t-1} + e_{1t} \\ z_t = a_{20} + a_{21}y_{t-1} + a_{22}z_{t-1} + e_{2t} \end{cases}$$

$$\varepsilon_t = Be_t \qquad B = \begin{pmatrix} 1 & b_{12} \\ b_{21} & 1 \end{pmatrix}.$$

The variance-covariance matrix of e_t is

$$\Sigma \equiv \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} = Ee_te_t' = B^{-1}E\varepsilon_t\varepsilon_t'B'^{-1}$$

$$= \frac{1}{(1-b_{12}b_{21})^2}\begin{pmatrix} \sigma_y^2 + b_{12}^2\sigma_z^2 & -(b_{21}\sigma_y^2 + b_{12}\sigma_z^2) \\ -(b_{21}\sigma_y^2 + b_{12}\sigma_z^2) & \sigma_z^2 + b_{21}^2\sigma_y^2 \end{pmatrix}.$$

VAR: Stationary

$$\begin{cases} y_t = a_{10} + a_{11} y_{t-1} + a_{12} z_{t-1} + e_{1t} \\ z_t = a_{20} + a_{21} y_{t-1} + a_{22} z_{t-1} + e_{2t} \end{cases}$$

$$\begin{cases} (1 - a_{11} L) y_t = a_{10} + a_{12} L z_t + e_{1t} \\ (1 - a_{22} L) z_t = a_{20} + a_{21} L y_t + e_{2t} \end{cases}$$

$$\boxed{\begin{aligned} y_t &= \frac{a_{10}(1 - a_{22}) + a_{12} a_{20} + (1 - a_{22} L) e_{1t} + a_{12} e_{2t-1}}{(1 - a_{11} L)(1 - a_{22} L) - a_{12} a_{21} L^2} \\ z_t &= \frac{a_{20}(1 - a_{11}) + a_{21} a_{10} + (1 - a_{11} L) e_{2t} + a_{21} e_{1t-1}}{(1 - a_{11} L)(1 - a_{22} L) - a_{12} a_{21} L^2} \end{aligned}}$$

which show that: 1) $\{y_t\}$ and $\{z_t\}$ have the same characteristic equation, and hence $\{y_t\}$ and $\{z_t\}$ **exhibit similar time paths**; 2) The stability condition for $\{y_t\}$ and $\{z_t\}$ requires that the roots of the polynomial equation $(1 - a_{11} L)(1 - a_{22} L) - a_{12} a_{21} L^2 - 0$ lie outside the unit circle, or, equivalently, the roots (the characteristic roots of the matrix A_1) of the following characteristic equation of A_1 :

$$\begin{vmatrix} \lambda - a_{11} & -a_{12} \\ -a_{21} & \lambda - a_{22} \end{vmatrix} = 0$$

lie inside the unit circle.

VAR: Stationary

Under the stationarity,

$$x_t = A_0 + A_1 x_{t-1} + e_t$$

$$= \mu + \sum_{i=0}^{\infty} A_1^i e_{t-i} \quad (\text{here } A_1^{n+1} \to 0 \text{ as } n \to \infty),$$

where $\mu = (I - A_1)^{-1} A_0$.

the variance-covariance matrix of x_t is

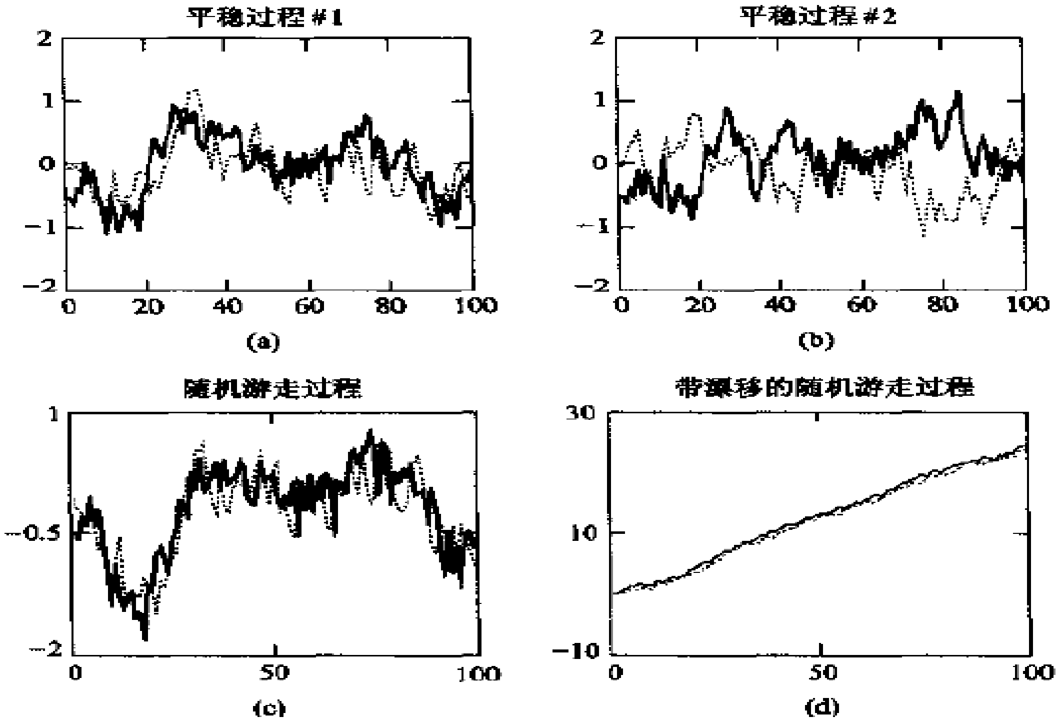
$$E(x_t - \mu)(x_t - \mu)' = \sum_{i,j=0}^{\infty} A_1^i E e_{t-i} e'_{t-j} \left(A_1^j\right)'$$

$$= \sum_{i=0}^{\infty} A_1^i E e_{t-i} e'_{t-i} \left(A_1^i\right)' = \sum_{i=0}^{\infty} A_1^i \Sigma \left(A_1^i\right)'.$$

VAR Process: Co-movement Pattern

1). $\begin{cases} y_t = 0.7y_{t-1} + 0.2z_{t-1} + e_{1t} \\ z_t = 0.2y_{t-1} + 0.7z_{t-1} + e_{2t} \end{cases}$

2). $\begin{cases} y_t = 0.5y_{t-1} - 0.2z_{t-1} + e_{1t} \\ z_t = -0.2y_{t-1} + 0.5z_{t-1} + e_{2t} \end{cases}$



3). $\begin{cases} y_t = 0.5y_{t-1} + 0.5z_{t-1} + e_{1t} \\ z_t = 0.5y_{t-1} + 0.5z_{t-1} + e_{2t} \end{cases}$

4). $\begin{cases} y_t = 0.5 + 0.5y_{t-1} + 0.5z_{t-1} + e_{1t} \\ z_t = 0.5y_{t-1} + 0.5z_{t-1} + e_{2t} \end{cases}$

VAR: Identification

- Estimating the structural VAR is inappropriate because of the endogeneity of z_t in the first equation and the endogeneity of y_t in the second equation.
- However, there is no such problem in estimating the reduced-form VAR model. We obtain nine estimates for the parameters in the reduced-form.
- Can recover the parameters in the original structural VAR model from those estimates in the reduced-form VAR model? No!

VAR: Identification Problem

- In structural VAR model, there are ten parameters to be determined, but in the reduced VAR we only have gotten 9 estimated parameters →Underidentification
- We have to restrict the original parameters to ensure that they can be solved by the relationship between the parameters in the two forms of the model

VAR: Identification

- **Choleski decomposition:** Impose a restriction on the primitive system: $b_{21} = 0$

which means that y_t does not have a contemporaneous effect on z_t.

$$\begin{cases} y_t = b_{10} - b_{12} z_t + \gamma_{11} y_{t-1} + \gamma_{12} z_{t-1} + \varepsilon_{yt} \\ z_t = b_{20} \qquad + \gamma_{21} y_{t-1} + \gamma_{22} z_{t-1} + \varepsilon_{zt} \end{cases} \qquad \begin{cases} y_t = a_{10} + a_{11} y_{t-1} + a_{12} z_{t-1} + e_{1t} \\ z_t = a_{20} + a_{21} y_{t-1} + a_{22} z_{t-1} + e_{2t} \end{cases}$$

$$e_t = \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix} = \begin{pmatrix} 1 & b_{12} \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{pmatrix} = \begin{pmatrix} \varepsilon_{yt} - b_{12}\varepsilon_{zt} \\ \varepsilon_{zt} \end{pmatrix}$$

(both ε_{yt} and ε_{zt} shocks affect the contemporaneous value of y_t,

but only ε_{zt} shock affects the contemporaneous value of z_t,

ε_{yt} does not affect e_{2t}. $\mathbf{z}_t$ **is "causally prior" to** $\mathbf{y}_t$)

$$\varepsilon_t = Be_t, \text{ i.e. } \begin{pmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{pmatrix} == \begin{pmatrix} 1 & b_{12} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix} = \begin{pmatrix} e_{1t} - b_{12} e_{2t} \\ e_{2t} \end{pmatrix}.$$

VAR: Identification 1

- **Choleski decomposition:** Impose a restriction on the primitive system: $b_{21} = 0$

$$A_0 = \begin{pmatrix} a_{10} \\ a_{20} \end{pmatrix} = \begin{pmatrix} 1 & b_{12} \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} b_{10} \\ b_{20} \end{pmatrix} = \begin{pmatrix} 1 & -b_{12} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} b_{10} \\ b_{20} \end{pmatrix} = \begin{pmatrix} b_{10} - b_{12}b_{20} \\ b_{20} \end{pmatrix}$$

$$A_1 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} 1 & b_{12} \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & -b_{12} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{pmatrix} = \begin{pmatrix} \gamma_{11} - b_{12}\gamma_{21} & \gamma_{12} - b_{12}\gamma_{22} \\ \gamma_{21} & \gamma_{22} \end{pmatrix}$$

$$\Sigma \equiv \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix} \equiv \begin{pmatrix} var(e_{1t}) & cov(e_{1t}, e_{2t}) \\ cov(e_{1t}, e_{2t}) & var(e_{2t}) \end{pmatrix} = \begin{pmatrix} \sigma_y^2 + b_{12}^2\sigma_z^2 & -b_{12}\sigma_z^2 \\ -b_{12}\sigma_z^2 & \sigma_z^2 \end{pmatrix}$$

which constitute nine equations with nine unknowns. The parameters in structural model can be exactly identified and recovered from the estimates of the reduced model. Here the ordering of y_t and z_t is important.

VAR: Identification 2

Another solution: set $b_{12} = 0$ (z_t does not have a contemporaneous effect on y_t). The argument is similar. In this case, $\mathbf{y}_t$ **is "causally prior" to** $\mathbf{z}_t$.

Note: We can restrict the parameters in the way the derivation above works well, but it is better that we use the restrictions which have some economic meanings in the structural VAR model.

- Other solution? See P295-298

  b12=a known constant;

  structural variance = a constant;

  b12=b21

VAR: Impulse response function

$$
\begin{aligned}
x_t &= A_0 + A_1 x_{t-1} + e_t = \mu + \sum_{i=0}^{\infty} A_1^i e_{t-i} \\
&= \mu + \frac{1}{1 - b_{12}b_{21}} \sum_{i=0}^{\infty} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^i \begin{pmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_{yt-i} \\ \varepsilon_{zt-i} \end{pmatrix} \\
&\equiv \mu + \sum_{i=0}^{\infty} \begin{pmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{pmatrix} \varepsilon_{t-i} \equiv \mu + \sum_{i=0}^{\infty} \phi_i \varepsilon_{t-i}.
\end{aligned}
$$

Here $\phi_i = A_1^i \phi_0$, i.e.

$$
\begin{pmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{pmatrix} = \frac{1}{1 - b_{12}b_{21}} \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^i \begin{pmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{pmatrix}
$$

are called **impulse response function**, which, in practice, are constructed from the estimated coefficients in the reduced model. For example, $\phi_{12}(1)$ is the instantaneous **impact (multiplier)** of a one-unit change in in ε_{zt} on y_t; $\phi_{12}(1)$ is the one-period response of a one-unit change in ε_{zt} on y_t; $\sum_{i=0}^{n} \phi_{12}(i)$ is the cumulated sum of the effects of ε_{zt} on the $\{y_t\}$ sequence after n periods:

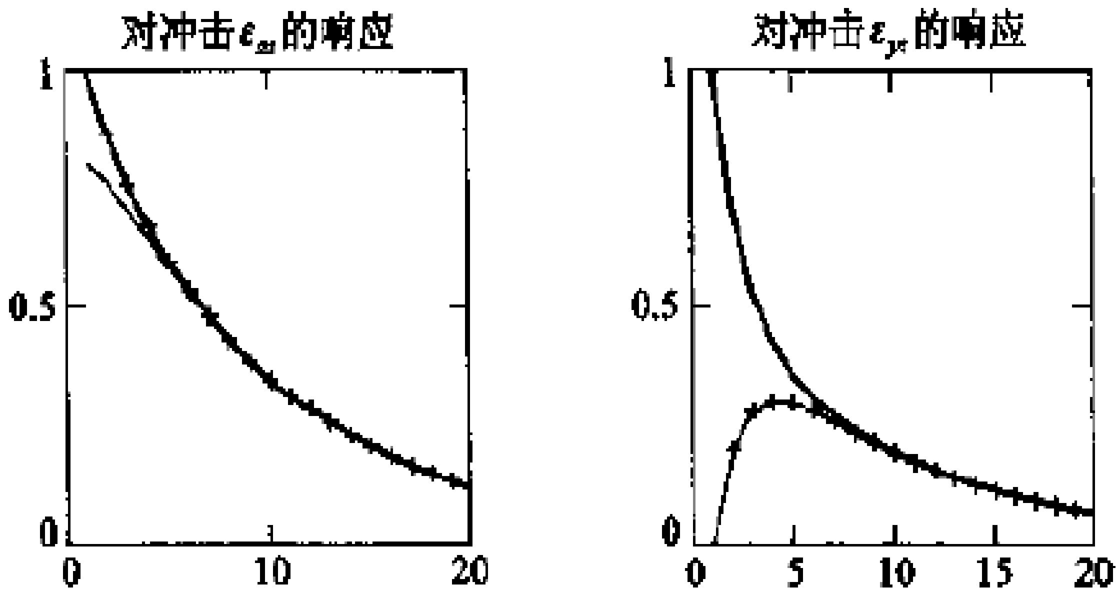
$$
\sum_{i=0}^{n} \phi_{12}(i) = \frac{\partial y_t}{\partial \varepsilon_{zt}} + \frac{\partial y_{t+1}}{\partial \varepsilon_{zt}} + \cdots + \frac{\partial y_{t+n}}{\partial \varepsilon_{zt}}.
$$

$\sum_{i=0}^{\infty} \phi_{12}(i)$ is the **long-run multiplier** of a one-unit change in in ε_{zt} on y_t, which is finite since $\{y_t\}$ and $\{z_t\}$ are assumed to be stationary.

Impulse response function
Example 1: $b_{21}=0$

$$\begin{pmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{pmatrix}$$

模型1: $\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} 0.7 & 0.2 \\ 0.2 & 0.7 \end{pmatrix} \cdot \begin{pmatrix} y_{t-1} \\ z_{t-1} \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}$ $\quad e_{1t}=0.8\varepsilon_{zt}+\varepsilon_{yt}$ 和 $e_{2t}=\varepsilon_{zt}$



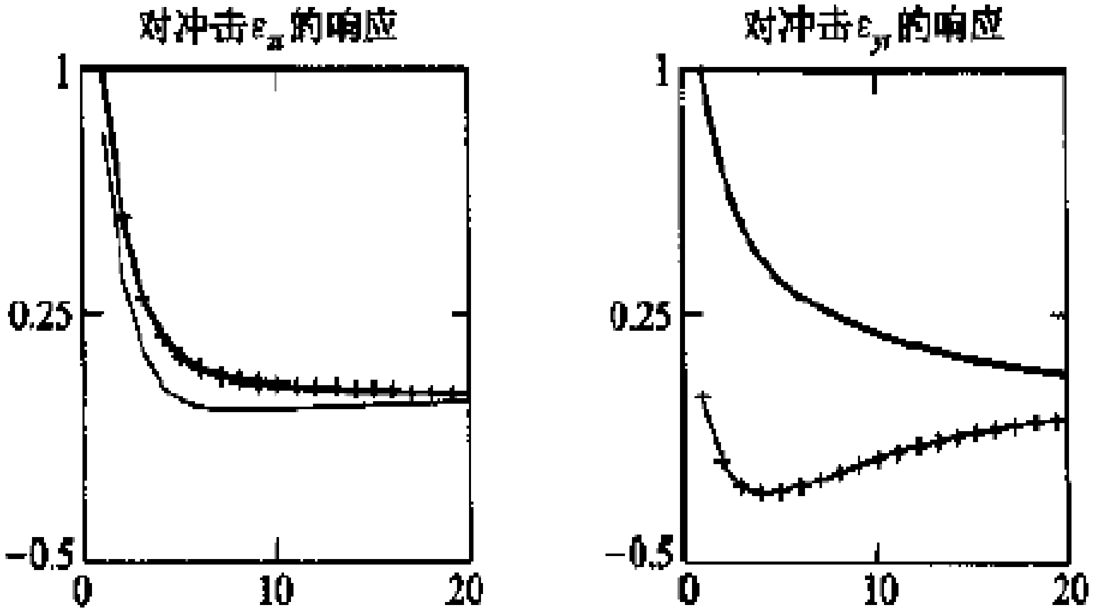
实线=(y_t)序列　　交叉的平行线=(z_t)序列

Impulse response function

Example 2: $b_{21}=0$

$$\begin{pmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{pmatrix}$$

模型 2: $\begin{pmatrix} y_t \\ z_t \end{pmatrix} = \begin{pmatrix} 0.7 & -0.2 \\ -0.2 & 0.7 \end{pmatrix} \cdot \begin{pmatrix} y_{t-1} \\ z_{t-1} \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \end{pmatrix}$　　$e_{1t}=0.8e_{zt}+\varepsilon_{yt}$ 和 $e_{2t}=\varepsilon_{zt}$



实线=(y_t)序列　　交叉的平行线−(z_t)序列

VAR: Forecast and Forecast Error

$$\boxed{x_t = A_0 + A_1 x_{t-1} + e_t,}$$

$$E_t x_{t+1} = A_0 + A_1 x_t \qquad x_{t+1} - E_t x_{t+1} = e_{t+1}.$$

$$E_t x_{t+2} = E_t \left(A_0 + A_1 (A_0 + A_1 x_t + e_{t+1}) + e_{t+2} \right) = (I + A_1)\, A_0 + A_1^2 x_t$$

$$x_{t+2} - E_t x_{t+2} = e_{t+2} + A_1 e_{t+1}$$

$$E_t x_{t+n} = \left(I + A_1 + \cdots + A_1^{n-1} \right) A_0 + A_1^n x_t$$

$$x_{t+n} - E_t x_{t+n} = e_{t+n} + A_1 e_{t+n-1} + \cdots + A_1^{n-1} e_{t+1}.$$

using the impulse response function $x_t = \mu + \sum_{i=0}^{\infty} \phi_i \varepsilon_{t-i}$ and $e_t = \phi_0 \varepsilon_t = B^{-1} \varepsilon_t,$

$$E_t x_{t+1} = \mu + \sum_{i=0}^{\infty} \phi_i E_t \varepsilon_{t+1-i} = \mu + \sum_{i=1}^{\infty} \phi_i \varepsilon_{t+1-i} \qquad x_{t+1} - E_t x_{t+1} = \phi_0 \varepsilon_{t+1}.$$

$$E_t x_{t+2} = \mu + \sum_{i=0}^{\infty} \phi_i E_t \varepsilon_{t+2-i} = \mu + \sum_{i=2}^{\infty} \phi_i \varepsilon_{t+2-i} \qquad x_{t+2} - E_t x_{t+2} = \phi_0 \varepsilon_{t+2} + \phi_1 \varepsilon_{t+1}.$$

$$E_t x_{t+n} = \mu + \sum_{i=0}^{\infty} \phi_i E_t \varepsilon_{t+n-i} = \mu + \sum_{i=n}^{\infty} \phi_i \varepsilon_{t+n-i} \qquad x_{t+1} - E_t x_{t+1} = \sum_{i=0}^{n-1} \phi_i \varepsilon_{t+n-i}.$$

VAR: Forecast Error Variance Decomposition

the n-step-ahead forecast error of y_{t+n} is

$$y_{t+n} - E_t y_{t+n} = \sum_{i=0}^{n-1} (\phi_{11}(i)\varepsilon_{yt+n-i} + \phi_{12}(i)\varepsilon_{zt+n-i})$$

the n-step-ahead forecast error variance of y_{t+n} is

$$\sigma_y^2(n) = \sum_{i=0}^{n-1} (\phi_{11}^2(i)\sigma_y^2 + \phi_{12}^2(i)\sigma_z^2) = \sigma_y^2 \sum_{i=0}^{n-1} \phi_{11}^2(i) + \sigma_z^2 \sum_{i=0}^{n-1} \phi_{12}^2(i).$$

The proportions of $\sigma_y^2(n)$ due to shocks in the $\{\varepsilon_{yt}\}$ and $\{\varepsilon_{zt}\}$ sequences are,

$$\frac{\sigma_y^2 \sum_{i=0}^{n-1} \phi_{11}^2(i)}{\sigma_y^2(n)} : \textit{the proportion due to its own shock}$$

$$\frac{\sigma_z^2 \sum_{i=0}^{n-1} \phi_{12}^2(i)}{\sigma_y^2(n)} : \textit{the proportion due to the shock of the other variable } z_t.$$

This is the **forecast error variance decomposition.** If ε_{zt} shocks explain none of the variance of y_t at all forecast horizons, $\{y_t\}$ is exogenous, and hence $\{y_t\}$ evolves independently of ε_{zt} shocks and $\{z_t\}$; If ε_{zt} shocks explain all of the forecast error variance of y_t at all forecast horizons, $\{y_t\}$ is entirely endogenous.

VAR: Forecast Error Variance Decomposition

Under the restriction $b_{21} = 0$, the original structural model is identified: $e_{1t} = \varepsilon_{yt} - b_{12}\varepsilon_{zt}$ and $e_{2t} = \varepsilon_{zt}$. All of the one-period forecast error variance of z_t is due to ε_{zt}, since, as $n = 1$, by (2),

$$\frac{\sigma_y^2 \sum_{i=0}^{n-1} \phi_{21}^2(i)}{\sigma_z^2(n)} = \frac{\sigma_y^2 \phi_{21}^2(0)}{\sigma_z^2(1)} = 0 \text{ or } \frac{\sigma_z^2 \sum_{i=0}^{n-1} \phi_{22}^2(i)}{\sigma_z^2(n)} = \frac{\sigma_z^2 \phi_{22}^2(0)}{\sigma_z^2(1)} = 1.$$

Similarly, under the restriction $b_{12} = 0$, since $e_{1t} = \varepsilon_{yt}$ and $e_{2t} = -b_{21}\varepsilon_{yt} + \varepsilon_{zt}$, all of the one-period forecast error variance of y_t is due to ε_{yt}.

n-equation VAR (the reduced form)

$$\begin{pmatrix} x_{1t} \\ x_{2t} \\ \vdots \\ x_{nt} \end{pmatrix} = \begin{pmatrix} A_{10} \\ A_{20} \\ \vdots \\ A_{n0} \end{pmatrix} + \begin{pmatrix} A_{11}(L) & A_{12}(L) & \cdots & A_{1n}(L) \\ A_{21}(L) & A_{22}(L) & \cdots & A_{2n}(L) \\ \vdots & \vdots & & \vdots \\ A_{n1}(L) & A_{n2}(L) & \cdots & A_{nn}(L) \end{pmatrix} \begin{pmatrix} x_{1t-1} \\ x_{2t-1} \\ \vdots \\ x_{nt-1} \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \\ \vdots \\ e_{nt} \end{pmatrix}$$

$A_{ij}(L) = a_{ij}(1) + a_{ij}(2)L \cdots + a_{ij}(p)L^{p-1}$ (the polynomial in the lag operator L).

Maximum Likelihood Estimation

Suppose $e_t \sim iidN(0, \Sigma)$, $t = 1, 2, \cdots, T$, and $x_{1-p}, \cdots, x_0$ are given constant vectors.

$$f(e_t) = (|2\pi\Sigma|)^{-1/2} \exp\left(-\frac{1}{2}e_t'\Sigma^{-1}e_t\right)$$

$$\log L(x_1, x_2, \cdots, x_T; A_0, A_1, \Sigma) = \frac{-nT}{2}\ln(2\pi) - \frac{T}{2}\log|\Sigma| - \frac{1}{2}\sum_{t=1}^{T} e_t'\Sigma^{-1}e_t$$

where $e_t = x_t - A_0 - A_1 x_{t-1}$.

$$\boxed{\log L^*(x_1, x_2, \cdots, x_T; A_0, A_1) = C - \frac{T}{2}\log\left|\sum_{t=1}^{T} e_t e_t'\right| = c_0 - \frac{T}{2}\log|\hat{\Sigma}|,}$$

$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T} e_t e_t',$

n-equation VAR (the reduced form)

$$\begin{pmatrix} x_{1t} \\ x_{2t} \\ \vdots \\ x_{nt} \end{pmatrix} = \begin{pmatrix} A_{10} \\ A_{20} \\ \vdots \\ A_{n0} \end{pmatrix} + \begin{pmatrix} A_{11}(L) & A_{12}(L) & \cdots & A_{1n}(L) \\ A_{21}(L) & A_{22}(L) & \cdots & A_{2n}(L) \\ \vdots & \vdots & & \vdots \\ A_{n1}(L) & A_{n2}(L) & \cdots & A_{nn}(L) \end{pmatrix} \begin{pmatrix} x_{1t-1} \\ x_{2t-1} \\ \vdots \\ x_{nt-1} \end{pmatrix} + \begin{pmatrix} e_{1t} \\ e_{2t} \\ \vdots \\ e_{nt} \end{pmatrix}$$

$A_{ij}(L) = a_{ij}(1) + a_{ij}(2)L \cdots + a_{ij}(p)L^{p-1}$ (the polynomial in the lag operator L).

How to select the lag length p? Begin with the longest plausible lag length and set the VAR model as **the Unrestricted Model;** Determine whether a shorter lag length is appropriate. Restrict the coefficients of $x'_{t-i}s$ for the lags between the longest lag length and this shorter lag length to be zero and obtain **the Restricted Model**. Then examine the significance of the null of these zero coefficients by **using the χ^2 statistic (likelihood ratio test):**

$$(T-c)\left(\log|\Sigma_r| - \log|\Sigma_u|\right),$$

where T is the sample size used in the estimation; c is the number of parameters estimated in each equation of the unrestricted model; $|\Sigma_u|$ is the determinant of the variance-covariance matrix of the residuals from the unrestricted VAR model; $|\Sigma_r|$ is the determinant of the variance-covariance matrix of the residuals from the restricted VAR model. The degree = the number of restrictions. Large value (greater than the critical value) of this sample statistic implies a rejection of the restriction; hence use the model with the longer lag length. **AIC and SBC**

VAR: Granger Causality

Granger Causality: How to test whether the lags of one variable enter into the equation for another variable in a VAR model?

$$x_j \textit{ does not } \textbf{Granger cause } x_i \Leftrightarrow \textit{all the coefficients of } A_{ij}(L) \textit{ equal zero.}$$

Thus, if $\{x_{jt}\}$ does not improve the forecasting performance of $\{x_{it}\}$, $\{x_{jt}\}$ does not Granger cause $\{x_{it}\}$. If all variables in the VAR model are stationary, conduct a standard F-test of the restriction

$$a_{ij}(1) = a_{ij}(2) = \cdots = a_{ij}(p) = 0.$$

Notice the difference between Granger causality and exogeneity: "$\{y_t\}$ Granger cause $\{z_t\}$" refers to the effects of **past values** of $\{y_t\}$ on the current value of z_t, and hence, Granger causality measures whether current and past values of y_t help to forecast future values of $\{z_t\}$; "z_t is exogenous in the equation of y_t" means that z_t is not affected by the **contemporaneous** value of y_t. Study the example

$$z_t = \bar{z} + \phi_{21}(0)\varepsilon_{yt} + \sum_{i=0}^{\infty} \phi_{22}(i)\varepsilon_{zt-i}.$$

Here $\{y_t\}$ does not Granger cause $\{z_t\}$, but z_t is endogenous in the equation for y_t.

VAR: Block-Causality

Block-causality test: How to determine whether p lags of one variable (say, w_t) enter the equations of any other variables (say y_t and z_t) in the system. Whether does w_t Granger cause y_t or z_t?

$$First, estimate\ the\ y_t\ and\ z_t\ equations \text{ using } lags\ of\ y_t, z_t\ and\ w_t \Rightarrow \Sigma_u;$$

$$Second, estimate\ the\ y_t\ and\ z_t\ equations\ only \text{ using } lags\ of\ y_t\ and\ z_t \Rightarrow \Sigma_r;$$

$$Then, construct\ the\ likelihood\ ratio\ statistic:$$

$$(T - 3p - 1)\left(\log|\Sigma_r| - \log|\Sigma_u|\right) \sim \chi^2(2p).$$

VAR: Tests with nonstationary variables

If the coefficient of interest can be written as a coefficient on a stationary variable, a t-test and F-test are appropriate, even though other variables are nonstationary.

1). Consider a two-variable VAR model: $\{y_t\} \sim I(1)$ and $\{z_t\} \sim I(0)$

$$y_t = a_{11}y_{t-1} + a_{12}y_{t-2} + b_{11}z_{t-1} + b_{12}z_{t-2} + \varepsilon_t.$$

$H_0: b_{11} = 0$ $\quad H_0: b_{12} = 0,$ $\quad H_0: b_{11} = b_{12} = 0,$

$H_0: a_{11} = 0$ $\quad H_0: a_{12} = 0$ $\quad \boxed{H_0: a_{11} = a_{12} = 0}$

2). Consider a two-variable VAR model:

$$y_t = a_{11}y_{t-1} + a_{12}y_{t-2} + b_{11} \quad \{y_t\} \sim I(1) \text{ and } \{z_t\} \sim I(1)$$

$$H_0: a_{12} = b_{12} = 0.$$

3). We may be able to test Granger causality between two nonstationary variables.

$$y_t = \gamma_1 y_{t-1} + a_{11}\Delta y_{t-1} + a_{12}\Delta y_{t-2} + b_{11}\Delta z_{t-1} + b_{12}\Delta z_{t-2} + c_{10}x_{t-1} + c_{11}\Delta x_{t-1} + c_{12}\Delta x_{t-2} + \varepsilon_t,$$

where $\{y_t\} \sim I(1)$, $\{z_t\} \sim I(1)$, $\{x_t\} \sim I(1)$

the test whether $\{z_t\}$ Granger causes $\{y_t\}$ is appropriate, but the test whether $\{x_t\}$ Granger causes $\{y_t\}$ is inappropriate.

Multivariate Structural VAR

$$\begin{pmatrix} 1 & b_{12} & \cdots & b_{1n} \\ b_{21} & 1 & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & 1 \end{pmatrix} \begin{pmatrix} x_{1t} \\ x_{2t} \\ \vdots \\ x_{nt} \end{pmatrix} = \begin{pmatrix} b_{10} \\ b_{20} \\ \vdots \\ b_{n0} \end{pmatrix} + \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1n} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2n} \\ \vdots & \vdots & & \vdots \\ \gamma_{n1} & \gamma_{n2} & \cdots & \gamma_{nn} \end{pmatrix} \begin{pmatrix} x_{1t-1} \\ x_{2t-1} \\ \vdots \\ x_{nt-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \vdots \\ \varepsilon_{nt} \end{pmatrix}$$

$$Bx_t = \Gamma_0 + \Gamma_1 x_{t-1} + \varepsilon_t.$$

The reduced-form VAR is $x_t = B^{-1}\Gamma_0 + B^{-1}\Gamma_1 x_{t-1} + B^{-1}\varepsilon_t$

$$= A_0 + A_1 x_{t-1} + e_t, \qquad A_0 = B^{-1}\Gamma_0,\ A_1 = B^{-1}\Gamma_1,\ e_t = B^{-1}\varepsilon_t$$

$$\Sigma \equiv Ee_te_t' = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{pmatrix} = EB^{-1}\varepsilon_t\varepsilon_t'B'^{-1} = B^{-1}E(\varepsilon_t\varepsilon_t')B'^{-1} \equiv B^{-1}\Sigma_\varepsilon B'^{-1}$$

$$\Sigma_\varepsilon \equiv E\varepsilon_t\varepsilon_t' = \begin{pmatrix} var(\varepsilon_1) & 0 & \cdots & 0 \\ 0 & var(\varepsilon_2) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & var(\varepsilon_n) \end{pmatrix}$$

$$\boxed{\hat{A}_0,\ \hat{A}_1,\ \hat{\Sigma},\ \hat{e}_t}$$

$$\boxed{\mathbf{B},\ \mathbf{\Gamma}_0,\ \mathbf{\Gamma}_1,\ \mathbf{\Sigma}_\varepsilon,\ \varepsilon_t}$$

Multivariate Structural VAR

$$\boxed{\mathbf{B},\ \mathbf{\Gamma}_0,\ \mathbf{\Gamma}_1,\ \mathbf{\Sigma}_\varepsilon,\ \varepsilon_t}$$

$$\boxed{\hat{A}_0,\ \hat{A}_1,\ \hat{\Sigma},\ \hat{e}_t}$$

$$Bx_t = \Gamma_0 + \Gamma_1 x_{t-1} + \varepsilon_t.$$

$$\begin{aligned} x_t &= B^{-1}\Gamma_0 + B^{-1}\Gamma_1 x_{t-1} + B^{-1}\varepsilon_t \\ &= A_0 + A_1 x_{t-1} + e_t, \end{aligned}$$

$$\Sigma \equiv Ee_t e_t' = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{pmatrix} \quad \text{(symmetric)}$$

$$\Sigma_\varepsilon \equiv E\varepsilon_t\varepsilon_t' = \begin{pmatrix} var(\varepsilon_1) & 0 & \cdots & 0 \\ 0 & var(\varepsilon_2) & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & var(\varepsilon_n) \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} 1 & b_{12} & \cdots & b_{1n} \\ b_{21} & 1 & \cdots & b_{2n} \\ \vdots & \vdots & & \vdots \\ b_{n1} & b_{n2} & \cdots & 1 \end{pmatrix}$$

$$\begin{aligned} & \textit{the number of unknowns} - \textit{the number of equations} \\ = \;& n^2 - \frac{1}{2}n(n+1) = \frac{1}{2}n(n-1) > 0 \end{aligned}$$

the structural model is not identified.

necessary to impose n(n-1)/2 restrictions on the structural VAR.

Multivariate Structural VAR

We only need to study $e = B^{-1}\varepsilon$ and decide the number of the restrictions.

For example, for a three-variable VAR, a Choleski decomposition is

$$\begin{pmatrix} e_{1t} \\ e_{2t} \\ e_{3t} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ c_{21} & 1 & 0 \\ c_{31} & c_{32} & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{pmatrix}.$$

An alternative is

$$\begin{pmatrix} e_{1t} \\ e_{2t} \\ e_{3t} \end{pmatrix} = \begin{pmatrix} 1 & 0 & c_{13} \\ c_{21} & 1 & 0 \\ c_{31} & 0 & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \end{pmatrix}.$$

Structural VAR: Overidentification

more than $\frac{1}{2}n(n-1)$ restrictions

Overidentification Test for an overidentified system:

Step 1 Estimate the unrestricted VAR: $x_t = A_0 + A_1 x_{t-1} + \cdots + A_p x_{t-p} + e_t$. Use the standard lag length and block causality tests to determine the form of the VAR.

Step 2 Calculate the unrestricted variance/covariance matrix Σ and $|\Sigma|$.

Step 3 Restrict B and/or Σ_ε by the overindentifying restrictions and use MLE to estimate the restricted VAR model respect to the free parameters in B and Σ_ε. This will lead to an estimate of the restricted variance/covariance matrix, denoted as Σ_R. Calculate $|\Sigma_R|$.

Step 4 Construct the test statistic $\boxed{\chi^2 = |\Sigma_R| - |\Sigma| \sim \chi^2(R)}$

where R is the number of restrictions exceeding $\frac{1}{2}n(n-1)$. If the calculated value of χ^2 exceeds that in a χ^2 table, the restrictions can be rejected.