TA section 8

May 29th, 2009 by Yang

1. Asymptotic Distribution of MLE of VAR Model

Consider the VAR(p) model

$$\underbrace{y_t}_{n\times 1} = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \varepsilon_t$$

where $\varepsilon_t \sim N(0, \Omega)$. The conditional distribution of y_t given $y_{t-1}, \cdots, y_{-p+1}$ is

$$y_t | y_{t-1}, \cdots, y_{-p+1} \sim N(c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p}, \Omega)$$

So we would like to maximize

$$L(\theta) = \prod_{t=1}^{T} f(y_t | y_{t-1}, \cdots, y_{-p+1}; \theta)$$

or equivalently,

$$\mathcal{L}(\theta) = \sum_{t=1}^{T} \log f(y_t | y_{t-1}, \cdots, y_{t-p}; \theta)$$

It turns out that the conditional MLE of the parameters is numerically the same with the OLS estimator. Denote

$$\underbrace{X_t}_{(np+1)\times 1} = \begin{pmatrix} 1 \\ y_{t-1} \\ \vdots \\ y_{t-p} \end{pmatrix} \text{ and } \underbrace{\Pi}_{(np+1)\times n} = \begin{pmatrix} c' \\ \phi_1' \\ \vdots \\ \phi_p' \end{pmatrix}$$

then the VAR(p) model can be written as

$$y_t = \Pi' X_t + \varepsilon_t$$

and

$$\widehat{\Pi}_{MLE} = \left(\sum_{t=1}^{T} X_t X_t' \right)^{-1} \sum_{t=1}^{T} X_t y_t'$$

or equivalently,

$$\widehat{\Pi}_{MLE}' = \sum_{t=1}^{T} y_t X_t' \left(\sum_{t=1}^{T} X_t X_t' \right)^{-1}$$

Also define $\widehat{\varepsilon}_t = y_t - \widehat{\Pi}_{MLE}' X_t$, then

$$\widehat{\Omega}_{MLE} = \frac{1}{T} \sum_{t=1}^{T} \widehat{\varepsilon}_t \widehat{\varepsilon}_t'$$

If we write

$$\underbrace{y}_{T\times n} = \begin{pmatrix} y_1' \\ \vdots \\ y_T' \end{pmatrix}, \quad \underbrace{X}_{T\times(np+1)} = \begin{pmatrix} X_1' \\ \vdots \\ X_T' \end{pmatrix} \text{ and } \underbrace{\widehat{\varepsilon}}_{T\times n} = \begin{pmatrix} \widehat{\varepsilon}_1' \\ \vdots \\ \widehat{\varepsilon}_T' \end{pmatrix}$$

then,

$$\widehat{\Pi}_{MLE} = (X'X)^{-1}X'y$$
$$\widehat{\Omega}_{MLE} = \frac{1}{T}\widehat{\varepsilon}'\widehat{\varepsilon}$$

Define

$$\pi := vec\Pi \text{ and } \omega := vech\Omega$$

then, we have the following asymptotic distribution.

$$\sqrt{T}\begin{pmatrix} \widehat{\pi}_T - \pi \\ \widehat{\omega}_T - \omega \end{pmatrix} \xrightarrow{d} N\left(0, \begin{bmatrix} \Omega \otimes Q^{-1} & 0 \\ 0 & \Sigma_{22} \end{bmatrix}\right)$$

where

$$Q := EX_tX_t'$$
$$asym.cov(\widehat{\sigma}_{ij}, \widehat{\sigma}_{ml}) = \sigma_{il}\sigma_{jm} + \sigma_{im}\sigma_{jl}$$

2. Application to AR(1)

Consider

$$y_t = c + \rho y_{t-1} + \varepsilon_t$$

From the asymptotic result,

$$\sqrt{T}\begin{pmatrix} \widehat{c} - c \\ \widehat{\rho} - \rho \end{pmatrix} \xrightarrow{d} N\left(0, \sigma^2 \begin{bmatrix} 1 & Ey_{t-1} \\ Ey_{t-1} & Ey_{t-1}^2 \end{bmatrix}^{-1}\right)$$

Note first that

$$\mu := Ey_{t-1} = Ey_t = \frac{c}{1-\rho}$$
$$var(y_t) = \frac{\sigma^2}{1-\rho^2}$$

Apply these to obtain

$$\sqrt{T}\begin{pmatrix} \widehat{c} - c \\ \widehat{\rho} - \rho \end{pmatrix} \xrightarrow{d} N\left(0, \begin{bmatrix} \sigma^2 + \mu^2(1-\rho^2) & -\mu(1-\rho^2) \\ -\mu(1-\rho^2) & 1-\rho^2 \end{bmatrix}\right)$$

and thus

$$\sqrt{T}(\widehat{\rho} - \rho) \xrightarrow{d} N(0, 1-\rho^2)$$

which depends only on ρ, but not on c and σ^2.

3. Unit Root Case

The above result enables us to guess that when $\rho = 1$,

$$\sqrt{T}(\widehat{\rho} - \rho) \xrightarrow{d} N(0, 0)$$

which is equivalent to

$$\sqrt{T}(\widehat{\rho} - \rho) = o_p(1)$$

In order to get an asymptotic distribution, we need to blow up $\widehat{\rho} - \rho$ with a series growing faster than $\sqrt{T}$. A more sophiscated statistical theory shows

$$T(\widehat{\rho} - \rho) = O_p(1)$$

and actually it converges to a certain distribution. Try some other exponents to verify this. If we use $\beta < 1$,

$$T^{\beta}(\widehat{\rho} - \rho) = o_p(1)$$

and if $\beta > 1$,

$$T^{\beta}(\widehat{\rho} - \rho) \text{ diverges}$$

4. Causality Test

Consider the VAR(2) model

$$\underbrace{y_t}_{2\times 1} = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t$$

where $\varepsilon_t \sim N(0, \Omega)$. Following the question 3 in Problem Set 8, write it as

$$y_t = c + \Phi^{(1)} y_{t-1} + \Phi^{(2)} y_{t-2} + \varepsilon_t$$

which reads as

$$\begin{pmatrix} y_{t1} \\ y_{t2} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \Phi_{11}^{(1)} & \Phi_{12}^{(1)} \\ \Phi_{21}^{(1)} & \Phi_{22}^{(1)} \end{pmatrix} \begin{pmatrix} y_{t-1,1} \\ y_{t-1,2} \end{pmatrix} + \begin{pmatrix} \Phi_{11}^{(2)} & \Phi_{12}^{(2)} \\ \Phi_{21}^{(2)} & \Phi_{22}^{(2)} \end{pmatrix} \begin{pmatrix} y_{t-2,1} \\ y_{t-2,2} \end{pmatrix} + \begin{pmatrix} \varepsilon_{t1} \\ \varepsilon_{t2} \end{pmatrix}$$

If $H_0 : \Phi_{12}^{(1)} = \Phi_{12}^{(2)} = 0$ is true,

$$y_{t1} = c_1 + \Phi_{11}^{(1)} y_{t-1,1} + \Phi_{11}^{(2)} y_{t-2,1} + \varepsilon_{t1}$$

so the past history of y_{t2} does not affect y_{t1} conditional on the past history of y_{t1}. Testing H_0 means to check if there is a causal effect of y_{t2} on y_{t1}. To do this, estimate first

$$\widehat{\Pi} = (X'X)^{-1} X'y = \left(\sum_{t=1}^{T} X_t X_t' \right)^{-1} \sum_{t=1}^{T} X_t y_t$$

and obtain

$$\widehat{\Omega} = \frac{1}{T}\widehat{\varepsilon}'\widehat{\varepsilon} = \frac{1}{T}\sum_{t=1}^{T}\widehat{\varepsilon}_t\widehat{\varepsilon}_t' \xrightarrow{p} \Omega$$
$$\widehat{Q} = \frac{1}{T}X'X = \frac{1}{T}\sum_{t=1}^{T}X_tX_t' \xrightarrow{p} EX_tX_t' = Q$$

Since

$$\sqrt{T}(\underbrace{vec\widehat{\Pi}}_{=:\widehat{\pi}} - vec\Pi) = \sqrt{T}\left[\begin{pmatrix} \widehat{c}_1 \\ \widehat{\Phi}_{11}^{(1)} \\ \widehat{\Phi}_{12}^{(1)} \\ \widehat{\Phi}_{11}^{(2)} \\ \widehat{\Phi}_{12}^{(2)} \\ \widehat{c}_2 \\ \widehat{\Phi}_{21}^{(1)} \\ \widehat{\Phi}_{22}^{(1)} \\ \widehat{\Phi}_{21}^{(2)} \\ \widehat{\Phi}_{22}^{(2)} \end{pmatrix} - \begin{pmatrix} c_1 \\ \Phi_{11}^{(1)} \\ \Phi_{12}^{(1)} \\ \Phi_{11}^{(2)} \\ \Phi_{12}^{(2)} \\ c_2 \\ \Phi_{21}^{(1)} \\ \Phi_{22}^{(1)} \\ \Phi_{21}^{(2)} \\ \Phi_{22}^{(2)} \end{pmatrix}\right] \xrightarrow{d} N(0, \Omega \otimes Q^{-1})$$

we have under the null hypothesis,

$$\sqrt{T}\left[\underbrace{\begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}}_{=:R}\begin{pmatrix} \widehat{c}_1 \\ \widehat{\Phi}_{11}^{(1)} \\ \widehat{\Phi}_{12}^{(1)} \\ \widehat{\Phi}_{11}^{(2)} \\ \widehat{\Phi}_{12}^{(2)} \\ \widehat{c}_2 \\ \widehat{\Phi}_{21}^{(1)} \\ \widehat{\Phi}_{22}^{(1)} \\ \widehat{\Phi}_{21}^{(2)} \\ \widehat{\Phi}_{22}^{(2)} \end{pmatrix} - \underbrace{\begin{pmatrix} 0 \\ 0 \end{pmatrix}}_{=:r}\right] \xrightarrow{d} N\left(0, R(\Omega \otimes Q^{-1})R'\right)$$

so the Wald statistic has the following asymptotic distribution.

$$W := T(R\widehat{\pi} - r)'\left[R(\widehat{\Omega} \otimes \widehat{Q}^{-1})R'\right]^{-1}(R\widehat{\pi} - r) \xrightarrow{d} \chi_2^2$$

5. Other Problems

(a) Nonstationarity of Unit Root Model

Consider the AR(1) model

$$y_t = \rho y_{t-1} + \varepsilon_t$$

where $\varepsilon_t \sim iid\ WN(0, \sigma^2)$ and $\sigma^2 > 0$. There cannot be a stationary causal solution y_t if $\rho = 1$.

Proof. Suppose there exists such a solution. The AR(1) equation together with $\rho = 1$ implies

$$y_t^2 = y_{t-1}^2 + 2y_{t-1}\varepsilon_t + \varepsilon_t^2$$

and taking expectation,

$$Ey_t^2 = Ey_{t-1}^2 + 2Ey_{t-1}\varepsilon_t + E\varepsilon_t^2$$

From stationarity, $Ey_t^2 = Ey_{t-1}^2$ and from causality, $Ey_{t-1}\varepsilon_t = 0$, so the above equation reads as

$$0 = \sigma^2$$

which is a contradiction.

(b) Optimal Linear Forecast of Stationary Time Series

Let $\alpha' X_t$ be the optimal linear forecast of y_{t+1}, where $X_t = \binom{1}{y_{t-1}}$. By the condition of the optimal linear forecast,

$$E\left[y_{t+1} - \alpha' \begin{pmatrix} 1 \\ y_{t-1} \end{pmatrix}\right] \begin{pmatrix} 1 & y_{t-1} \end{pmatrix} = 0$$

Since

$$E \begin{pmatrix} 1 \\ y_{t-1} \end{pmatrix} \begin{pmatrix} 1 & y_{t-1} \end{pmatrix} = E \begin{pmatrix} 1 & y_{t-1} \\ y_{t-1} & y_{t-1}^2 \end{pmatrix} = \begin{pmatrix} 1 & \mu \\ \mu & \gamma_0 + \mu^2 \end{pmatrix}$$

and

$$Ey_{t+1} \begin{pmatrix} 1 & y_{t-1} \end{pmatrix} = E \begin{pmatrix} y_{t+1} & y_{t+1}y_{t-1} \end{pmatrix} = \begin{pmatrix} \mu & \gamma_2 + \mu^2 \end{pmatrix}$$

we have

$$\begin{aligned} \alpha' &= \begin{pmatrix} \mu & \gamma_2 + \mu^2 \end{pmatrix} \begin{pmatrix} 1 & \mu \\ \mu & \gamma_0 + \mu^2 \end{pmatrix}^{-1} \\ &= \frac{1}{\gamma_0} \begin{pmatrix} \mu & \gamma_2 + \mu^2 \end{pmatrix} \begin{pmatrix} \gamma_0 + \mu^2 & -\mu \\ -\mu & 1 \end{pmatrix} \\ &= \begin{pmatrix} \frac{(\gamma_0 - \gamma_2)\mu}{\gamma_0} & \frac{\gamma_2}{\gamma_0} \end{pmatrix} \end{aligned}$$

So the optimal linear forecast of y_{t+1} based on 1 and y_{t-1} is

$$\widehat{P}(y_{t+1}|1, y_{t-1}) = \alpha' X_t = \frac{(\gamma_0 - \gamma_2)\mu}{\gamma_0} + \frac{\gamma_2}{\gamma_0} y_{t-1}$$

Apply this to the AR(1) model

$$y_t = c + \rho y_{t-1} + \varepsilon_t$$

Note that

$$\gamma_0 = \frac{\sigma^2}{1 - \rho^2} \text{ and } \gamma_h = \rho^h \gamma_0$$

Therefore,

$$\widehat{P}(y_{t+1}|1, y_{t-1}) = \alpha' X_t = \mu(1 - \rho^2) + \rho^2 y_{t-1}$$

(c) Subcollection of AR(1) Processes

Consider the AR(1) model

$$y_t = c + \rho y_{t-1} + \varepsilon_t$$

where $\varepsilon_t \sim iid\ WN(0, \sigma^2)$. Now suppose we observe $z_t := y_{2t}$. We claim z_t is also an AR(1) process. To prove this, we need to show that z_t satisfies the first order linear difference equation

$$z_t = d + \phi z_{t-1} + u_t$$

with $u_t \sim iid\ WN(0, \omega^2)$. Note first that

$$\begin{aligned} z_t = y_{2t} &= c + \rho y_{2t-1} + \varepsilon_{2t} \\ &= \underbrace{c + \rho c}_{=:d} + \underbrace{\rho^2}_{=:\phi} y_{2t-2} + \underbrace{\rho \varepsilon_{2t-1} + \varepsilon_{2t}}_{=:u_t} \\ &= d + \phi z_{t-1} + u_t \end{aligned}$$

Finally verify that u_t is an iid WN process by checking

$$\begin{aligned} Eu_t &= E\varepsilon_{2t} + \rho E\varepsilon_{2t-1} = 0 \\ \omega^2 := Eu_t^2 &= E\varepsilon_{2t}^2 + 2\rho E\varepsilon_{2t}\varepsilon_{2t-1} + \rho^2 E\varepsilon_{2t-1}^2 = (1 + \rho^2)\sigma^2 \\ Eu_t u_{t-h} &= E\varepsilon_{2t}\varepsilon_{2t-2h} + E\varepsilon_{2t}\varepsilon_{2t-2h-1} + E\varepsilon_{2t-1}\varepsilon_{2t-2h} + E\varepsilon_{2t-1}\varepsilon_{2t-2h-1} = 0 \end{aligned}$$

for any t and $h \neq 0$.