Parameter Estimation and Central Limit Theorem for AR(p) Model

Ching-Kang Ing

National Tsing Hua University

cking@stat.nthu.edu.tw

Overview

1. CLT for $\hat{\phi}_p$
2. Estimate σ^2

CLT for $\hat{\phi}_p$

- Remember we have showed that, for AR(1) models,

$$\sqrt{n}(\hat{\phi}_1 - \phi) \xrightarrow{d} N(0, \gamma_0^{-1}\sigma^2).$$

  We are now in a position to generalize this result to AR(p) models.

- Recall that

$$(\hat{\phi}_p - \phi_p) = (\mathbf{X}_p^\top \mathbf{X}_p)^{-1}\mathbf{X}_p^\top \mathbf{y} - \phi_p,$$

  where

$$\mathbf{X}_p = \begin{pmatrix} Z_p & \cdots & Z_1 \\ \vdots & \ddots & \vdots \\ Z_{n-1} & \cdots & Z_{n-p} \end{pmatrix} \quad \text{and} \quad \mathbf{Y} = \begin{pmatrix} Z_{p+1} \\ \vdots \\ Z_n \end{pmatrix}.$$

- Since $\phi_p = (\mathbf{X}_p^\top \mathbf{X}_p)^{-1}\mathbf{X}_p'\mathbf{X}_p\phi_p$, it can be shown that

$$\hat{\phi}_p - \phi_p = (n^{-1}\mathbf{X}_p^\top \mathbf{X}_p)^{-1}\frac{1}{n}\sum_{j=p}^{n-1}\mathbf{Z}_j(p)\epsilon_{j+1},$$

where $\mathbf{Z}_j(p) = (Z_j, \cdots, Z_{j-p+1})^\top$.
- We have argued that $n^{-1}\mathbf{X}_p^\top \mathbf{X}_p \xrightarrow{p} \mathbf{R}_p$ and hence $(n^{-1}\mathbf{X}_p^\top \mathbf{X}_p)^{-1} \xrightarrow{p} \mathbf{R}_p^{-1}$, implying

$$\sqrt{n}(\hat{\phi}_p - \phi_p) \dot{\sim} \mathbf{R}_p^{-1}\frac{1}{\sqrt{n}}\sum_{j=p}^{n-1}\mathbf{Z}_j(p)\epsilon_{j+1}.$$

Therefore, if we can show that

$$\frac{1}{\sqrt{n}}\sum_{j=p}^{n-1}\mathbf{Z}_j(p)\varepsilon_{j+1} \xrightarrow{d} N(\mathbf{0}, \mathbf{R}_p\sigma^2), \tag{1}$$

then

$$\sqrt{n}(\hat{\phi}_p - \phi_p) \xrightarrow{d} N(\mathbf{0}, \mathbf{R}_p^{-1}\sigma^2).$$

- To show (1), we need some background knowledge.

Definition

A sequence of random vectors, $\{\mathbf{w}_n\}$, is said to converge to a random vector, $\mathbf{w}$, in distribution if

$$\Pr(\mathbf{w}_n \leq \mathbf{c}) \to \Pr(\mathbf{w} \leq \mathbf{c}) \equiv F(\mathbf{c}) \quad \text{as } n \to \infty$$

for all continuous points of $F(\cdot)$. The convergence of $\mathbf{w}_n$ to $\mathbf{w}$ in distribution is denoted by

$$\mathbf{w}_n \xrightarrow{d} \mathbf{w}.$$

Cramér-Wold device

$\mathbf{w}_n \xrightarrow{d} \mathbf{w} \Leftrightarrow \mathbf{a}^\top \mathbf{w}_n \xrightarrow{d} \mathbf{a}^\top \mathbf{w} \quad \text{for all } \|\mathbf{a}\| = 1.$

- By making use of the Cramér-wold device, it suffices for (1) to show that

$$\frac{1}{\sqrt{n}} \sum_{j=p}^{n-1} [\mathbf{a}^\top \mathbf{Z}_j(p)] \epsilon_{j+1} \xrightarrow{d} N(0, \mathbf{a}^\top \mathbf{R}_p \mathbf{a} \sigma^2) \tag{2}$$

for any $\mathbf{a} \in \mathbf{R}^p$ with $\|\mathbf{a}\| = 1$. Now, (2) can be easily proved in the same manner as the used in AR(1) models.

Estimate σ^2

- We can use

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{j=p}^{n-1}\left(Z_{j+1} - \hat{\phi}_p^\top \mathbf{Z_j}(p)\right)^2$$

to estimate σ^2.

- Note that

$$\begin{aligned}\hat{\sigma}^2 &= \frac{1}{n}\sum_{j=p}^{n-1}\left(\epsilon_{j+1} - (\hat{\phi}_p - \phi_p)^\top \mathbf{Z_j}(p)\right)^2 \\ &= \frac{1}{n}\sum_{j=p}^{n-1}\epsilon_{j+1}^2 - \left(\frac{1}{n}\sum_{j=p}^{n-1}\mathbf{Z_j}(p)\epsilon_{j+1}\right)^\top \tilde{\mathbf{R}}_p^{-1}\left(\frac{1}{n}\sum_{j=p}^{n-1}\mathbf{Z_j}(p)\epsilon_{j+1}\right),\end{aligned}$$

where $\tilde{\mathbf{R}}_p = \frac{1}{n}\mathbf{X}_p^\top\mathbf{X}_p = \frac{1}{n}\sum_{j=p}^{n-1}\mathbf{Z_j}(p)\mathbf{Z_j}^\top(p)$.

- Since we've argued that $\tilde{\mathbf{R}}_p^{-1} \xrightarrow{p} \mathbf{R}_p^{-1}$ (which is positive definite) and $n^{-1/2} \sum_{j=p}^{n-1} \mathbf{Z}_j(p)\epsilon_{j+1} \xrightarrow{d} N(\mathbf{0}, \mathbf{R}_p\sigma^2)$, it follows that

$$\frac{(n^{-1/2} \sum_{j=p}^{n-1} \mathbf{Z}_j(p)\varepsilon_{j+1})^\top \tilde{\mathbf{R}}_p^{-1} (n^{-1/2} \sum_{j=p}^{n-1} \mathbf{Z}_j(p)\varepsilon_{j+1})}{\sigma^2} \xrightarrow{d} \chi^2(p),$$

and hence

$$\Big(\frac{1}{n}\sum_{j=p}^{n-1} \mathbf{Z}_j(p)\epsilon_{j+1}\Big)^\top \tilde{\mathbf{R}}_p^{-1} \Big(\frac{1}{n}\sum_{j=p}^{n-1} \mathbf{Z}_j(p)\epsilon_{j+1}\Big) = O_p(n^{-1}) = o_p(1).$$

- Moreover, by the law of large number for the sums of i.i.d random variables, we have

$$\frac{1}{n}\sum_{j=p}^{n-1} \varepsilon_{j+1}^2 \xrightarrow{pr} E(\varepsilon_i) = \sigma^2.$$

- Combining these facts yields

$$\hat{\sigma}^2 \xrightarrow{p} \sigma^2.$$

Moreover, by the CLT for the sum of i.i.d r.v.s, one gets

$$\sqrt{n}(\hat{\sigma}^2 - \sigma^2) \dot{\sim} \frac{1}{\sqrt{n}} \sum_{j=p}^{n-1} (\varepsilon_{j+1}^2 - \sigma^2) \xrightarrow{d} N\left(0, E(\varepsilon_i^2 - \sigma^2)^2\right).$$

- **Remark.** We don't need "normality" in proving the consistency and the CLT for $\hat{\sigma}^2$.