Lecture 12: Estimation (continued)

Bus 41910, Time Series Analysis, Mr. R. Tsay

Asymptotic properties of MLE: For a stationary and invertible ARMA model

$$\phi(B)Z_t = \theta(B)a_t$$

where $\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$ and $\theta(B) = 1 - \theta_1 B - \cdots - \theta_q B^q$ are polynomials with no common factors and $\{a_t\}$ is a sequence of *iid* Gaussian random variables with mean zero and variance σ_a^2, the MLE (both conditional and exact) of the parameters ϕ's and θ's are (a) consistent and (b) asymptotically normal. More precisely, let $\boldsymbol{\beta} = (\phi_1, \cdots, \phi_p, \theta_1, \cdots, \theta_q)'$, and $\hat{\boldsymbol{\beta}}$ the MLE of $\boldsymbol{\beta}$. Then,

$$\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \sim N(\boldsymbol{O}, \boldsymbol{V}_\beta) \quad \text{as} \quad n \to \infty$$

where n is the sample size and $\boldsymbol{V}_\beta$ is a $(p+q) \times (p+q)$ covariance matrix defined by

$$\boldsymbol{V}_\beta = \sigma_a^2 \begin{bmatrix} E(\boldsymbol{u}_t \boldsymbol{u}_t') & E(\boldsymbol{u}_t \boldsymbol{v}_t') \\ E(\boldsymbol{v}_t \boldsymbol{u}_t') & E(\boldsymbol{v}_t \boldsymbol{v}_t') \end{bmatrix}^{-1}$$

where $\boldsymbol{u}_t = (u_t, u_{t-1}, \cdots, u_{t-p+1})'$ and $\boldsymbol{v}_t = (v_t, v_{t-1}, \cdots, v_{t-q+1})'$ with u_t and v_t satisfying

$$\phi(B)u_t = a_t, \quad \theta(B)v_t = -a_t.$$

The asymptotic normality of the MLE follows basically the usual argument as that of the *iid* case with CLT replaced by certain functional central limit theorem. You may consult standard time series textbooks for details, e.g. Box and Jenkins (1976) and Brockwell and Davis (1991, Ch. 8 & 10). Here we shall discuss the reason supporting the asymptotic variance $\boldsymbol{V}_\beta$.

Recall that the log-likelihood function of Z_t is approximately

$$\ell_n(\boldsymbol{\beta}) \propto -\frac{n}{2}\ln\sigma_a^2 - \frac{1}{2\sigma_a^2}\sum_{t=1}^{n} a_t^2$$

where $a_t = Z_t - \phi_1 Z_{t-1} - \cdots - \phi_p Z_{t-p} + \theta_1 a_{t-1} + \cdots + \theta_q a_{t-q}$. The asymptotic covariance matrix of the MLE of $\boldsymbol{\beta}$ is then the inverse of the expected Fisher information matrix

$$E\left[-\frac{\partial^2 \ell_n(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}'}\right] = E\left[(\frac{\partial \ell_n(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}})(\frac{\partial \ell_n(\boldsymbol{\beta})}{\partial\boldsymbol{\beta}})'\right].$$

From the log-likelihood function, it is important to consider the derivatives of a_t with respect to $\boldsymbol{\beta}$ in order to compute the Fisher information matrix. It can also be shown that the MLE of $\boldsymbol{\beta}$ is asymptotically uncorrelated with that of σ_a^2 so that we can work on $\boldsymbol{\beta}$ and σ_a^2 separately.

Let

$$-\frac{\partial a_t}{\partial \phi_i} = u_{t-i}, \quad -\frac{\partial a_t}{\partial \theta_j} = v_{t-j}.$$

From $a_t = Z_t - \phi_1 Z_{t-1} - \cdots - \phi_p Z_{t-p} + \theta_1 a_{t-1} + \cdots + \theta_q a_{t-q}$, we have

$$u_{t-i} = Z_{t-i} + \theta_1 u_{t-1-i} + \cdots + \theta_q u_{t-q-i}$$

and

$$v_{t-j} = \theta_1 v_{t-1-j} + \cdots + \theta_q v_{t-q-j} - a_{t-j}.$$

In other words, we have

$$\theta(B) u_t = Z_t, \quad \text{and} \quad \theta(B) v_t = -a_t.$$

Next, since $Z_t = \frac{\theta(B)}{\phi(B)} a_t$, we further obtain

$$\theta(B) u_t = \frac{\theta(B)}{\phi(B)} a_t$$

so that

$$\phi(B) u_t = a_t.$$

From the stationarity and invertibility of Z_t, both u_t and v_t are stationary processes. Therefore, the Fisher information matrix is

$$\boldsymbol{V}_\beta^{-1} = \frac{n}{\sigma_a^2} E\left[\begin{pmatrix} \boldsymbol{u}_t \\ \boldsymbol{v}_t \end{pmatrix} (\boldsymbol{u}_t', \boldsymbol{v}_t')\right]$$

where $\boldsymbol{u}_t$ and $\boldsymbol{v}_t$ are defined as above.
In practice, the Fisher information matrix is evaluated by substituting $\boldsymbol{\beta}$ by $\hat{\boldsymbol{\beta}}$.

Some special cases: In what follows, we consider results of some simple ARMA models.
AR(p) Model: $\phi(B) Z_t = a_t$. For this model, the derived process u_t is an AR(p) process also. Therefore,

$$\boldsymbol{V}_\beta^{-1} = \frac{n}{\sigma_a^2} \boldsymbol{\Gamma}_p$$

where $\boldsymbol{\Gamma}_p$ is the covariance matrix of $(u_t, \cdots, u_{t-p+1})'$, or equivalently, the covariance matrix of $(Z_t, \cdots, Z_{t-p+1})'$. Consequently, we have

- AR(1): $\hat{\phi}_1 \sim N(\phi_1, \frac{1-\phi_1^2}{n})$.

- AR(2): $\hat{\boldsymbol{\phi}} \sim N(\boldsymbol{\phi}, \boldsymbol{V})$ where $\boldsymbol{\phi} = (\phi_1, \phi_2)'$ and

$$\boldsymbol{V} = \frac{1}{n}\begin{bmatrix} 1-\phi_2^2 & -\phi_1(1+\phi_2) \\ -\phi_1(1+\phi_2) & 1-\phi_2^2 \end{bmatrix}.$$

MA(q) Model: $Z_t = \theta(B)a_t$. Since the derived process v_t satisfies $\theta(B)v_t = -a_t$, which is an AR(q) model, we obtain the same results as those of pure AR models.

- MA(1): $\hat{\theta}_1 \sim N(\theta_1, \frac{1-\theta_1^2}{n})$.
- MA(2): $\hat{\theta} \sim N(\boldsymbol{\theta}, V)$ where $\boldsymbol{\theta} = (\theta_1, \theta_2)'$ and

$$\boldsymbol{V} = \frac{1}{n}\begin{bmatrix} 1-\theta_2^2 & -\theta_1(1+\theta_2) \\ -\theta_1(1+\theta_2) & 1-\theta_2^2 \end{bmatrix}.$$

Mixed ARMA(1,1) Model: $Z_t - \phi Z_{t-1} = a_t - \theta a_{t-1}$. For this process, the asymptotic covariance matrix of MLE of $\boldsymbol{\beta} = (\phi, \theta)'$ is

$$\boldsymbol{V} = \frac{1}{n}\frac{1-\phi\theta}{(\phi-\theta)^2}\begin{bmatrix} (1-\phi^2)(1-\phi\theta) & (1-\phi^2)(1-\theta^2) \\ (1-\phi^2)(1-\theta^2) & (1-\theta^2)(1-\phi\theta) \end{bmatrix}.$$