Time Series Analysis
Lecture 4 ARIMA modelling

Muyi Li

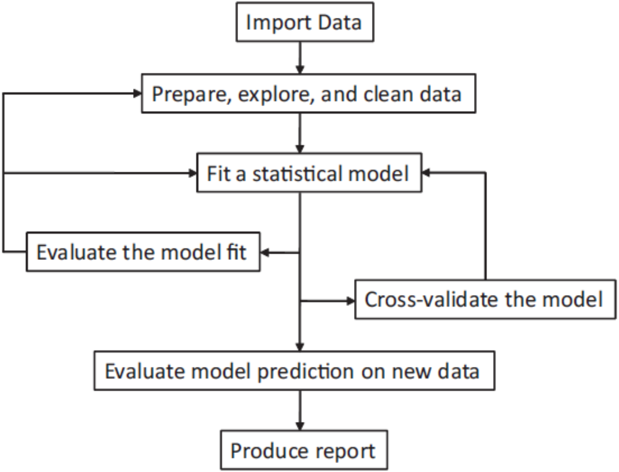
Wang ya'nan Institute for Studies in Economics (WISE)
Department of Statistics, School of Economics
Xiamen University

Outline

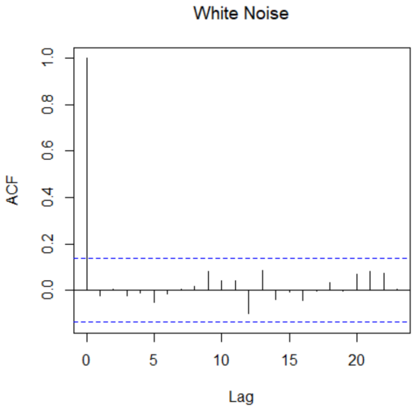
- Model Steps
- Model Specification
- Model Estimation
- Model Checking
- Model Selection

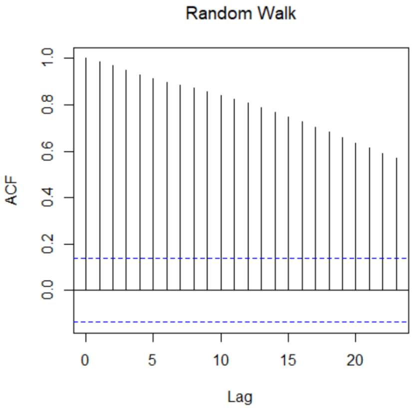
Modeling Steps

All models are wrong, but some are useful—George, E.P.Box

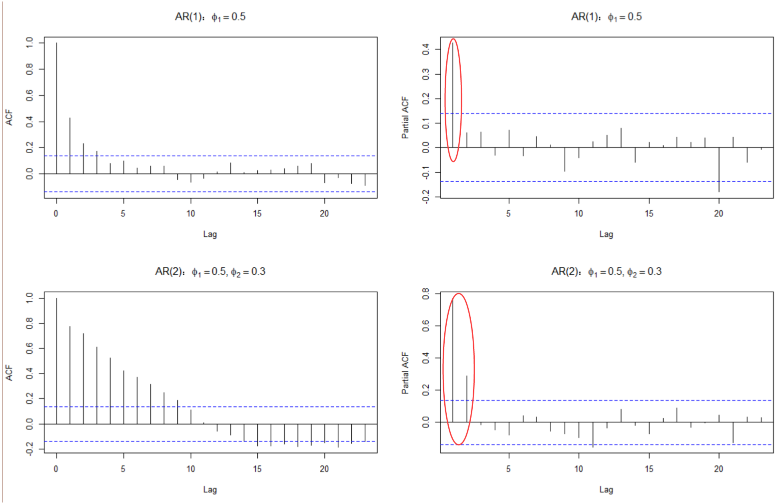


Illustrations of Simulated Time Series

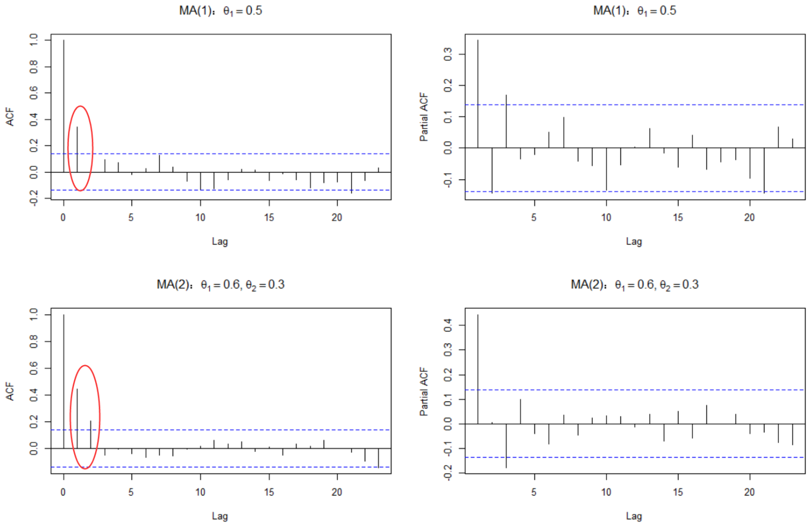




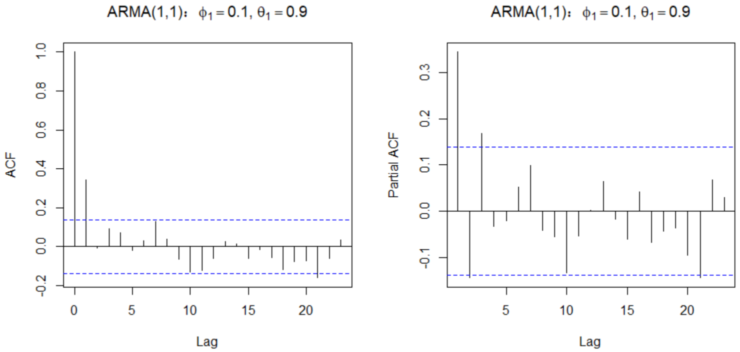
Illustrations of Simulated Time Series



Illustrations of Simulated Time Series



Illustrations of Simulated Time Series



Model Specification

1. If the sample ACF $\hat{\rho}_k$ are close to 1 for all small k and decay slowly, difference the data first.存在非平稳结构
2. For the sample ACF $\hat{\rho}_k$, we have

$$\sqrt{T}\hat{\rho}_k \xrightarrow{d} N(0, 1 + 2\sum_{j=1}^{k-1} \rho_j^2),$$

   hence if it is bounded by $\frac{1.96}{\sqrt{T}}\left[1 + 2\sum_{j=1}^{k-1} \hat{\rho}_j^2\right]^{1/2}$ for any $k > q$, we can try an MA(q) model to the data.
3. For the sample PACF $\hat{\pi}_k$,

$$\sqrt{T}\hat{\pi}_k \xrightarrow{d} N(0, 1),$$

   hence, if it is bounded by $1.96/\sqrt{T}$ for any $k > p$, we can try an AR(p) model to the data.
4. In general, the correct (p, q) is not unique. We can select the best one by AIC/BIC.

Ordinary Least Square of AR(p) Model

- **Consider an AR(p) model,**

$$y_t = \phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + e_t.$$

The parameter space $\theta = (\phi', \sigma^2), \phi = (\phi_0, \phi_1, \cdots, \phi_p)'$.

- **The OLS estimates of coefficients:**

$$\begin{aligned}(\widehat{\phi}_0, \widehat{\phi}_1, ..., \widehat{\phi}_p) &= argmin \sum_{t=p+1}^{T} (y_t - \phi_0 - \phi_1 y_{t-1} - \cdots - \phi_p y_{t-p})^2 \\ &= argmin S_C(\phi_0, \phi_1, ..., \phi_p)\end{aligned}$$

OLS of AR(p)

- Let $x_t = (1, y_{t-1}, y_{t-2}, ..., y_{t-p})'$ and $\phi = (\phi_0, \phi_1, ..., \phi_p)'$. Then,

$$\begin{aligned}\widehat{\phi}_{OLS} &= \left(\sum_{t=p+1}^{T} x_t x_t'\right)^{-1} \sum_{t=p+1}^{T} x_t y_t \\ &= \phi + \left(\sum_{t=p+1}^{T} x_t x_t'\right)^{-1} \sum_{t=p+1}^{T} x_t e_t.\end{aligned}$$

- $$\widehat{\sigma}_e^2 = \frac{1}{T-p} S_C(\widehat{\phi}_0, \widehat{\phi}_1, ..., \widehat{\phi}_p).$$

MLE for AR(p) Model

- Consider the following $AR(p)$ process:

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \varepsilon_t,$$

  where ε_t is i.i.d $N(0, \sigma^2)$.
- The parameter vector is $\theta = \{\phi_0, , \phi_1, \phi_2, ..., \phi_p, \sigma^2\}$.
- For Y_t, $t > p$, the conditional distribution is

$$(Y_t | Y_{t-1}, Y_{t-2}, ..., Y_{t-p}) \sim N(\phi_0 + \phi_1 y_{t-1} + \phi_2 y_{t-2} ... + \phi_p y_{t-p}, \sigma^2)$$

- the conditional density is

$$\begin{aligned} & f_{Y_t|Y_{t-1},Y_{t-2},...,Y_{t-p}}(y_t|y_{t-1}, y_{t-2}, ...y_{t-p}, \theta) \\ = \quad & \frac{1}{\sqrt{2\pi}\sigma} \exp(\frac{-(y_t - \phi_0 - \phi_1 y_{t-1} - \phi_2 y_{t-2} ... - \phi_p y_{t-p})^2}{2\sigma^2}). \end{aligned}$$

MLE for AR(p) process

- Assume $\mathbf{y}_p = (y_1, y_2, \cdots, y_p)$.
- Conditional MLE: assume $\mathbf{y}_p$ to be deterministic.
- The conditional likelihood function

$$L(\theta) = \sum_{t=p+1}^{T} \log(f_{Y_t|Y_{t-1},Y_{t-2},\ldots,Y_{t-p}}(y_t|y_{t-1}, y_{t-2}, \ldots y_{t-p}, \theta))$$

$$= -(T-p)/2 \log(2\pi\sigma^2) - \sum_{t=p+1}^{T} \frac{(y_t - \phi_0 - \phi_1 y_{t-1} - \ldots - \phi_p y_{t-p})^2}{2\sigma^2}.$$

MLE of AR(p) model

- Unconditional MLE: assume $\mathbf{y}_p$ to be random vector
- the log likelihood function is

$$\begin{aligned} L(\theta) &= \log(f_{Y_T, Y_{T-1}, Y_{T-2}, \ldots, Y_1}(y_T, y_{T-1}, \ldots, y_1, \theta)) \\ &= -T/2 \log(2\pi) - T/2 \log(\sigma^2) + \\ &\quad \frac{1}{2\sigma^2} \log|V_p^{-1}| - \frac{1}{2\sigma^2}(\mathbf{y}_P - \mu_p)' V_p^{-1}(\mathbf{y}_P - \mu_p) \\ &\quad - \sum_{t=p+1}^{T} \frac{(y_t - \phi_0 - \phi_1 y_{t-1} - \phi_2 y_{t-2} \ldots - \phi_p y_{t-p})^2}{2\sigma^2} \end{aligned}$$

Conditional MLE for AR(1)

Consider AR(1) model: $y_t = \phi_0 + \phi_1 y_{t-1} + e_t$, $e_t \sim i.i.d.(0, \sigma^2)$.

- Assume y_1 be given as deterministic.

$$y_2 | y_1 = y_1 \sim N(\phi_0 + \phi_1 y_1, \sigma^2)$$

- The conditional density function is

$$f(y_2 | y_1, \phi) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(y_2 - \phi_0 - \phi_1 y_1)^2}{2\sigma^2}\right).$$

- The conditional joint density is

$$f(y_T, y_{T-1}, ..., y_2 | y_1, \theta) = \Pi_{t=2}^{T} f(y_t | y_{t-1}, \theta)$$

Conditional MLE for AR(1)

- The conditional log likelihood function

$$
\begin{aligned}
L(\theta) &= \sum_{t=2}^{T} \log(f(y_t|y_{t-1}, \theta)) \\
&= -(\frac{T-1}{2}) \log\left(2\pi\sigma^2\right) - \sum_{t=2}^{T} \frac{(y_t - \phi_0 - \phi y_{t-1})^2}{2\sigma^2}.
\end{aligned}
$$

- Theconditional MLE for θ is defined as

$$
\widehat{\theta} = \arg\max_{\theta} L(\theta).
$$

Conditional MLE for AR(1)

For $\widehat{\sigma}^2$, the first order condition is

$$-\frac{T-1}{2\sigma^2}+\sum_{t=2}^{T}\frac{(y_t-\phi_0-\phi y_{t-1})^2}{2\sigma^4}=0.$$

and the solution is

$$\widehat{\sigma}^2=\sum_{t=2}^{T}\frac{(y_t-\widehat{\phi_0}-\widehat{\phi}y_{t-1})^2}{T-1}.$$

Model Checking

Testing for residual autocorrelations

- $$H_0 : \rho_1 = \rho_2 = \cdots = 0 \leftrightarrow H_1 : \exists \quad j, \text{ such that} \rho_j \neq 0.$$
- Joint significance test of the first m residual autocorrelations.
- Ljung-Box typed test statistics:

$$T_m = n(n+2) \sum_{k=1}^{m} \frac{\rho_k^2(\hat{e})}{n-k} \sim \chi^2_{m-K}.$$

where

$$\rho_k(\hat{e}) = \frac{\sum_{t=k+1}^{n} \hat{e}_t \hat{e}_{t-k}}{\sum_{t=k+1}^{n} \hat{e}_t^2}$$

and K is the number of unknown parameters.

Testing for Normality

- The skewness and kurtosis of $\hat{e}_t$ can be calculated as

$$\widehat{SK}_{\hat{e}} = \frac{\hat{m}_3}{\sqrt{\hat{m}_2^3}}, \quad \text{and} \quad \widehat{K}_{\hat{e}} = \frac{\hat{m}_4}{\hat{m}_2^2}.$$

  where $\hat{m}_j = \frac{1}{n}\sum_{t=1}^{n} \hat{e}_t^j$,

- Under the null hypothesis of normality,

$$\sqrt{n/6} \cdot \widehat{SK}_{\hat{e}} \sim \mathcal{N}(0,1), \sqrt{n/24} \cdot (\widehat{K}_{\hat{e}} - 3) \sim \mathcal{N}(0,1).$$

- A joint test for normality (Jarque and Bera, 1987):

$$\text{JB} = \frac{n}{6}\widehat{SK}_{\hat{e}}^2 + \frac{n}{24}(\widehat{K}_{\hat{e}} - 3)^2 \sim \chi_2^2.$$

Model Diagnostic Checking

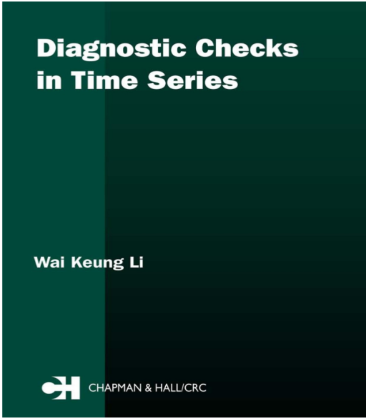


Figure 1: Li,WaiKeung.(2004). Diagnostic Checks in Time Series. Chapman& Hall/CRC

Model Selection: Order Determination

- Akaike information criterion (AIC):$-2L(\widehat{\theta}) + 2k$
- Bayesian information criterion (BIC):$-2L(\widehat{\theta}) + \ln(T)k$
- Compared to the AIC, the BIC counteracts the overfitting tendency of the AIC.
- Other information criterions.