Identification and Diagnostics

Short overview

(1) Summary

(2) Model Identification (Selection)
  - ACF, PACF
  - Information Criteria
  - From general to specific

(3) Diagnostic Checking

Short Summary/Overview

• In previous lectures we have covered how to estimate and make inference in ARMA(p,q) models under the assumption that we know p and q. However, in practice, these parameters have to be estimated as well

There are several approaches to select lags p and q

- Using ACF and PACF;
- Using Information Criteria (most common in practice).
- Testing model from general to specific;

(2) Model Identification (Selection)

ACF and PACF

- The **main objective** of this part is to choose the values of p and q such that the ARMA(p,q) model fit the data best. Since in the most cases there are no a priory reasons to choose a particular model spesification, we must look at the data to determine which ARMA(p,q) model seems appropriate.

Using ACF and PACF.

One way to identify the parameters p and q is to recognize the structure of the theoretical ACF and PACF from the structure of the sample ACF and PACF.

Approach 1: ACF and PACF

Process	Autocorrelations	Partial Autocorrelations
MA(q)	$\begin{cases} \rho_k \neq 0, \ k \leq q \\ \rho_k = 0, \ k > q \end{cases}$	$\begin{cases} \text{Exponential decay} \\ \text{Oscillations possible} \\ P_k \underset{k \to \infty}{\to} 0 \end{cases}$
AR(p)	$\begin{cases} \text{Exponential decay} \\ \rho_k \underset{k \to \infty}{\to} 0 \\ \rho_k = \phi_1 \rho_{k-1} + \ldots + \phi_p \rho_{k-p} \end{cases}$	$\begin{cases} P_k \neq 0, \ k \leq p \\ P_k = 0, \ k > p \end{cases}$
ARMA(p,q)	$\begin{cases} \text{Irregular for } k\text{=}1,\ldots,p \\ \rho_k \underset{k \to \infty}{\to} 0 \end{cases}$	$P_k \underset{k \to \infty}{\to} 0$

ACF and PACF

Summary:

• In most practical situations, it is hard to identify in which of the three previous cases we have.

• By choosing an ARMA model we want to fit the data accurately and also to have a parsimonious model (to avoid overfitting).

• Additional tools are required to check goodness of fit and the number of parameters

Information Criteria for Model Selection

There are three classical Information Criteria to select model:

- Akaike's Information Criterion (AIC)
- Schwarz's Bayesian Information Criterion (BIC)
- Hannan-Quin Information Criterion (HQ)

Information Criteria for Model Selection

- Suppose that we have K alternative models: $M_1, M_2, ..., M_K$, where index of the model reflect number of parameters (i.e., model M_{10} has 10 parameters). Then the general form of the information criteria can be formulated as

$$IC(k) = \ln(\widehat{\sigma}_k^2) + k\frac{C(T)}{T},$$

where T is the sample size (number of observations); $\widehat{\sigma}_k^2$ is the estimated variance of the residuals in model M_k; $C(T)$ is the "penalty" term that varies depending on the selected information criterion.

Information Criteria for Model Selection

The AIC for ARMA(p,q) model $[\mathbf{C}(\mathbf{T}) = \mathbf{2}]$

$$AIC(k) = \ln(\widehat{\sigma}_k^2) + (p+q)\frac{2}{T},$$

For BIC for ARMA(p,q) $[\mathbf{C}(\mathbf{T}) = \ln(\mathbf{T})]$

$$IC(k) = \ln(\widehat{\sigma}_k^2) + (p+q)\frac{\ln(T)}{T},$$

For HQ for ARMA(p,q) $[\mathbf{C}(\mathbf{T}) = \mathbf{2}\ln\ln(\mathbf{T})]$

$$IC(k) = \ln(\widehat{\sigma}_k^2) + (p+q)\frac{2\ln\ln(T)}{T},$$

Information Criteria for Model Selection

Suppose we have a family of the AR models to chose from: AR(1), AR(2),..., AR(p). Then we select the one with smallest value of information criterion, i.e.,

$$\widehat{p} = \operatorname{argmin}\{IC(k)\ |k = 1, ..., p\}.$$

Theorem

Let y_t be a stationary AR(p) process. Suppose that the maximum order of the model we chose from is $p_{max} > p$ and we select the $\widehat{p}$ that minimizes the information criterion

$$IC(k) = \ln(\widehat{\sigma}_k^2) + k\frac{C(T)}{T},\ \textit{for } k = 1, ..., p_{max}.$$

Then $\widehat{p}$ is consistent iff $C(T) \to \infty$ and $C(T)/T \to 0$ when $T \to \infty$.

Remark

- *We can conclude from the theorem that AIC gives not consistent results, while BIC and HQ are consistent.*

- *It can be shown that AIC tends to overselect number of lags. Is it a serious problem?*

Short Summary Box-Jenkins procedure

- **From specific to general.** Box and Jenkins (1975) propouse a procedure to build an ARMA model in order to approximate sufficiently well the GDP starting from an observed time series. Their procedure is made up of the following steps:

  1. Preliminary analysis of the data;
  2. Identification of ARMA(p,q) model;
  3. Estimation of the parameters of the ARMA model;
  4. Diagnostic checks (is the estimated model adequate?)
     (a) If not go back to step 2;
     (b) If yes proceed to forecasting.

- Both information criteria or ACF/PACF approaches belong to "**From specific to general**" methodology.

From General to Specific

- **From general to specific.**

From general to specific methodology works **via testing**. For this we need to estimate the most general model that make sense for a given data set and then test weather AR and MA components are significant. Insignificant component should be removed. These procedure is repeated until only significant components are left.

Diagnostic Checking

• The final step of ARMA model building is diagnostic checking, i.e., the implementation of some checks on model adequacy.

• A most common check is residual analysis which can be realized through:

(a) graphical analysis: either on $(t, \widehat{\varepsilon}_t)$ or $(\widehat{\varepsilon}_t, \widehat{\varepsilon}_{t+1})$;

(b) Sample ACF of the residuals: testing the significance of each $\widehat{\rho}_k$

(c) Portmanteau statistics: assessing the overall acceptability of residuals autocorrelation with the Ljung-Box statistic (distributed as χ^2_m):

$$Q_m = T(T+2)\sum_{k=1}^{m} \frac{1}{T-k}\widehat{\rho}_k^2$$