Chapter 7
Parameter estimation, Diagnostic Checking and Model Selection

Real data: $y_1, y_2, \cdots, y_n$.

New data: $Z_1, Z_2, \cdots, Z_n$ follow a stationary and invertible **ARMA** model.

Let $\dot{Z}_t = Z_t - \mu$.

Suppose that $\dot{Z}_t$ follows **ARMA** model:

$$\begin{aligned}\dot{Z}_t &= \phi_1 \dot{Z}_{t-1} + \phi_2 \dot{Z}_{t-2} + \cdots + \phi_p \dot{Z}_{t-p} \\ &\quad + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}\,,\end{aligned}$$

where $a_t \sim i.i.d. N(0, \sigma_a^2)$. We need to find

p, q, $\phi_1, \cdots, \phi_p$, $\theta_1, \cdots, \theta_q$, μ and σ_a^2, which are called unknown parameters.

How to find?

Procedure:

Step 1. Determine (p, q) by the method in Chapter 6. If it is not easy to find p and q, you can try some different (p, q).

Step 2. Estimate $\phi_1, \cdots, \phi_p,\ \theta_1, \cdots, \theta_q,\ \mu$ and σ_a^2.

Step 3. Checking whether or not (p, q) is correct.

If it is not correct, try some different (p, q) and then go to **Step 2**.

Even if it is correct, we still need to try some different (p, q) in practice.

Step 4. In general, the correct (p, q) is not unique. We can select the best one by **AIC** and **BIC** criteria.

The above procedure is called building a model, fit a model, model fitting, or modeling.

Section 7.1 The method of moments.

Model: $\dot{Z}_t = Z_t - \mu$,

$\dot{Z}_t = \phi_1 \dot{Z}_{t-1} + \phi_2 \dot{Z}_{t-2} + \cdots + \phi_p \dot{Z}_{t-p} + a_t$.

Yule-Walker equations:

$$\begin{aligned}
\rho_1 &= \phi_1 + \phi_2\rho_1 + \phi_3\rho_2 + \cdots + \phi_p\rho_{p-1} \\
\rho_2 &= \phi_1\rho_1 + \phi_2 + \phi_3\rho_1 + \cdots + \phi_p\rho_{p-2} \\
\rho_3 &= \phi_1\rho_2 + \phi_2\rho_1 + \phi_3 + \cdots + \phi_p\rho_{p-3} \\
&\vdots \\
\rho_p &= \phi_1\rho_{p-1} + \phi_2\rho_{p-2} + \cdots + \phi_{p-1}\rho_1 + \phi_p
\end{aligned}$$

Note that ρ_k can be estimated by $\widehat{\rho}_k$. Solve the above equations:

$$\begin{bmatrix} \widehat{\phi}_1 \\ \widehat{\phi}_2 \\ \widehat{\phi}_3 \\ \vdots \\ \widehat{\phi}_p \end{bmatrix} = \begin{bmatrix} 1 & \widehat{\rho}_1 & \widehat{\rho}_2 & \cdots & \widehat{\rho}_{p-2} & \widehat{\rho}_{p-1} \\ \widehat{\rho}_1 & 1 & \widehat{\rho}_1 & \cdots & \widehat{\rho}_{p-3} & \widehat{\rho}_{p-2} \\ \widehat{\rho}_2 & \widehat{\rho}_1 & 1 & \cdots & \widehat{\rho}_{p-4} & \widehat{\rho}_{p-3} \\ \vdots & & & & & \vdots \\ \widehat{\rho}_{p-1} & \widehat{\rho}_{p-2} & \widehat{\rho}_{p-3} & \cdots & \widehat{\rho}_1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} \widehat{\rho}_1 \\ \widehat{\rho}_2 \\ \widehat{\rho}_3 \\ \vdots \\ \widehat{\rho}_p \end{bmatrix}$$

$\widehat{\phi}_i$, $i = 1, 2, \cdots, p$, are called the moment or Yule-Walker estimators of ϕ_i, $i = 1, 2, \cdots, p$. The moment estimator of σ_a^2 is

$$\widehat{\sigma}_a^2 = \widehat{\gamma}_0(1 - \widehat{\phi}_1\widehat{\rho}_1 - \widehat{\phi}_2\widehat{\rho}_2 - \cdots - \widehat{\phi}_p\widehat{\rho}_p).$$

Example 7.1. Model:

$$(Z_t - \mu) = \phi_1(Z_{t-1} - \mu) + a_t.$$

Given Z_t, $t = 1, \cdots, n$, find the Yule-Walker estimators of μ, ϕ_1 and σ_a^2.

The drawbacks of Yule-Walker estimation:

(1). It is not robust if the a_t is not normal.

(2). It is not easy to estimate the **MA**(q) or **ARMA**(p, q) model ($q \neq 0$).

Section 7.2.
Maximum likelihood (ML) method

Section 7.2.1
Conditional ML estimation

Model: $\dot{Z}_t = Z_t - \mu$,

$$\begin{aligned}\dot{Z}_t &= \phi_1 \dot{Z}_{t-1} + \phi_2 \dot{Z}_{t-1} + \cdots + \phi_p \dot{Z}_{t-p} \\ &\quad + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q} ,\end{aligned}$$

where $a_t \sim i.i.d. N(0, \sigma_a^2)$.

Denote $\phi = (\phi_1, \cdots, \phi_p)'$ and $\theta = (\theta_1, \cdots, \theta_q)'$.

Given $Z_1, Z_2, \cdots, Z_n$,

the joint density $f(Z_n, Z_{n-1}, \cdots, Z_1)$ is the likelihood function.

If the initial values $Z_* = (Z_{1-p}, Z_{2-p}, \cdots, Z_0)'$, $a_* = (a_{1-q}, a_{2-q}, \cdots, a_0)'$ are used,

$f(Z_n, Z_{n-1}, \cdots, Z_1 | Z_*, a_*)$ is called the conditional likelihood function.

$$f(Z_n, Z_{n-1}, \cdots, Z_1 | Z_*, a_*) = \prod_{i=1}^{n} f(Z_i | Z_{i-1}, \cdots, Z_1, Z_*$$

$$\begin{aligned}
\dot Z_1 &= \phi_1 \dot Z_0 + \phi_2 \dot Z_{-1} + \cdots + \phi_p \dot Z_{1-p} \\
&\quad + a_1 - \theta_1 a_0 - \cdots - \theta_q a_{1-q}\,, \\
&\quad a_1 \sim i.i.d. N(0, \sigma_a^2).
\end{aligned}$$

$\dot Z_1 | (\dot Z_*, a_*) \sim N(M_1, \sigma_a^2)$, where

$$\begin{aligned}
M_1 &= \phi_1 \dot Z_0 + \phi_2 \dot Z_{-1} + \cdots + \phi_p \dot Z_{1-p} \\
&\quad - \theta_1 a_0 - \cdots - \theta_q a_{1-q}\,,
\end{aligned}$$

$$f(\dot Z_1 | Z_*, a_*) = \frac{1}{\sqrt{2\pi\sigma_a^2}} e^{-\frac{(\dot Z_1 - M_1)^2}{2\,\sigma_a^2}}.$$

$$\begin{aligned}
a_1(\phi, \mu, \theta) &\equiv \dot Z_1 - M_1 \\
&= \dot Z_1 - \phi_1 \dot Z_0 - \phi_2 \dot Z_{-1} - \cdots - \phi_p \dot Z_{1-p} \\
&\quad + \theta_1 a_0 + \cdots + \theta_q a_{1-q}. \\
\dot Z_2 &= \phi_1 \dot Z_1 + \phi_2 \dot Z_0 + \cdots + \phi_p \dot Z_{2-p} \\
&\quad + a_2 - \theta_1 a_1 - \cdots - \theta_q a_{2-q}\,, \\
&\quad a_2 \sim i.i.d. N(0, \sigma_a^2).
\end{aligned}$$

$\dot Z_2 | (\dot Z_1, \dot Z_*, a_*) \sim N(M_2, \sigma_a^2)$, where

$$\begin{aligned}
M_2 &= \phi_1 \dot Z_1 + \phi_2 \dot Z_0 + \cdots + \phi_p \dot Z_{2-p} \\
&\quad - \theta_1 a_1(\phi, \mu, \theta) - \cdots - \theta_q a_{2-q}\,.
\end{aligned}$$

$$f(\dot Z_2 | \dot Z_1, Z_*, a_*) = \frac{1}{\sqrt{2\pi\sigma_a^2}} e^{-\frac{(\dot Z_2 - M_2)^2}{2\,\sigma_a^2}}.$$

$$\begin{aligned} a_2(\phi,\mu,\theta) &\equiv \dot{Z}_2 - M_2 \\ &= \dot{Z}_2 - \phi_1\dot{Z}_1 - \phi_2\dot{Z}_0 - \cdots - \phi_p\dot{Z}_{2-p} \\ &\quad +\theta_1 a_1(\phi,\mu,\theta) + \cdots + \theta_q a_{2-q}. \end{aligned}$$

$\dot{Z}_n|(\dot{Z}_{n-1},\cdots,\dot{Z}_1,\dot{Z}_*,a_*) \sim N(M_n,\sigma_a^2)$, where

$$\begin{aligned} M_n &= \phi_1\dot{Z}_{n-1} + \phi_2\dot{Z}_{n-2} + \cdots + \phi_p\dot{Z}_{n-p} \\ &\quad -\theta_1 a_{n-1}(\phi,\mu,\theta) - \cdots - \theta_q a_{n-q}(\phi,\mu,\theta)\,. \end{aligned}$$

$$f(\dot{Z}_n|\dot{Z}_{n-1},\dot{Z}_1,Z_*,a_*) = \frac{1}{\sqrt{2\pi\sigma_a^2}} e^{-\frac{(\dot{Z}_n - M_n)^2}{2\,\sigma_a^2}}.$$

$\vdots$

$$\begin{aligned} a_n(\phi,\mu,\theta) &\equiv \dot{Z}_n - M_n \\ &= \dot{Z}_n - \phi_1\dot{Z}_{n-1} - \phi_2\dot{Z}_{n-2} - \cdots - \phi_p\dot{Z}_{n-p} \\ &\quad +\theta_1 a_{n-1}(\phi,\mu,\theta) + \cdots + \theta_q a_{n-q}(\phi,\mu,\theta) \end{aligned}$$

Let

$$\begin{aligned} L_*(\phi,\mu,\theta,\sigma_a^2) &= f(Z_n,Z_{n-1},\cdots,Z_1|Z_*,a_*). \\ L_*(\phi,\mu,\theta,\sigma_a^2) &= \prod_{i=1}^{n}(2\pi\sigma_a^2)^{-1/2} e^{-\frac{(\dot{Z}_i - M_i)^2}{2\,\sigma_a^2}} \\ \ln L_*(\phi,\mu,\theta,\sigma_a^2) &= -\frac{n}{2}\ln(2\pi\sigma_a^2) - \frac{S_*(\phi,\mu,\theta)}{2\,\sigma_a^2}. \end{aligned}$$

where $S_*(\phi,\mu,\theta) = \sum_{t=1}^{n} a_t^2(\phi,\mu,\theta)$.

$\ln L_*(\phi, \mu, \theta, \sigma_a^2)$ is called the conditional log-likelihood function.

The maximizer of $\ln L_*(\phi, \mu, \theta, \sigma_a^2)$ is called the conditional maximum likelihood estimator (CMLE) of $(\phi, \mu, \theta, \sigma_a^2)$, denoted by $(\widehat{\phi}, \widehat{\mu}, \widehat{\theta}, \widehat{\sigma}_a^2)$.

$(\widehat{\phi}, \widehat{\mu}, \widehat{\theta})$ is also the minimizer of $S_*(\phi, \mu, \theta)$. So, $(\widehat{\phi}, \widehat{\mu}, \widehat{\theta})$ is called conditional least squares estimator (CLSE) of (ϕ, μ, θ).

In practice, we first find $(\widehat{\phi}, \widehat{\mu}, \widehat{\theta})$ by minimizing:

$$S_*(\phi, \mu, \theta) = \sum_{t=1}^{n} a_t^2(\phi, \mu, \theta)$$

Then calculate $\widehat{\sigma}_a^2$ by

$$\widehat{\sigma}_a^2 = \frac{S_*(\widehat{\phi}, \widehat{\mu}, \widehat{\theta})}{n - p - q - 1}.$$

How to minimize $S_*(\phi, \mu, \theta)$?

Example 7.2 Model:

$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \cdots + \phi_p Z_{t-p} + a_t$
where $a_t \sim i.i.d. N(0, \sigma_a^2)$.

Solution:

$$\begin{aligned} a_t(\phi) &= Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} - \cdots - \phi_p Z_{t-p} \\ &= Z_t - \widetilde{Z}'_{t-1}\phi\,, \end{aligned}$$

where $\widetilde{Z}_{t-1} = (Z_{t-1}, Z_{t-2}, \cdots, Z_{t-p})'$.

$$S_*(\phi) = \sum_{t=1}^{n} a_t^2(\phi)$$

$$\begin{aligned} \frac{\partial S_*(\phi)}{\partial \phi} &= 2\sum_{t=1}^{n} \frac{\partial a_t(\phi)}{\partial \phi} a_t(\phi) \\ &= -\sum_{t=1}^{n} \widetilde{Z}_{t-1} Z_t + \sum_{t=1}^{n} \widetilde{Z}_{t-1} \widetilde{Z}'_{t-1}\phi \end{aligned}$$

Note that $\frac{\partial S_*(\phi)}{\partial \phi}\big|_{\phi=\widehat{\phi}} = 0$, we have

$$\widehat{\phi} = \left(\sum_{t=1}^{n} \widetilde{Z}_{t-1}\widetilde{Z}'_{t-1}\right)^{-1} \left(\sum_{t=1}^{n} \widetilde{Z}_{t-1} Z_t\right).$$

Since $\frac{\partial^2 S_*(\phi)}{\partial\phi\partial\phi'} = \sum_{t=1}^{n} \widetilde{Z}_{t-1}\widetilde{Z}'_{t-1} > 0$, $\widehat{\phi}$ is the CMLE.

$$\widehat{\sigma}_a^2 = \frac{1}{n-p}\sum_{t=1}^{n}\left[Z_t - \widetilde{Z}'_{t-1}\widehat{\phi}\right]^2.$$

Remark:

(1). $\widehat{\phi}$ for **AR(p)** model is equivalent to Yule-Walker estimator and is called ordinary least squares estimator (**OLS**) in Section 7.4.

(2). for the **ARMA**(p,q) model with $q \neq 0$, how to minimize $S_*(\phi, \mu, \theta)$ will be given in Section 7.3.

(3). In the **CMLE** or **CLSE**, we simply take the initial values: $Z_* = 0$ or $a_* = 0$, or $Z_* = \widehat{\mu}$ and $a_* = 0$. These initial values do not affect on the estimators.

For example, in the **MA**(1) model:

$$\begin{aligned} a_{100}(\theta) &= Z_{100} + \theta a_{99}(\theta) \\ &= Z_{100} + \theta Z_{99} + \theta^2 Z_{98} + \cdots + \theta^{100} Z_0 + \cdots \end{aligned}$$

Example 7.3.

exchange Rates: TEN/USA (1970-2000).

7.2.2. Unconditional ML Estimation

Model: $\dot{Z}_t = Z_t - \mu$,

$$\begin{aligned}\dot{Z}_t &= \phi_1 \dot{Z}_{t-1} + \phi_2 \dot{Z}_{t-1} + \cdots + \phi_p \dot{Z}_{t-p} \\ &\quad + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q} ,\end{aligned}$$

where $a_t \sim i.i.d.N(0, \sigma_a^2)$.

Given $Z = (Z_n, Z_{n-1}, \cdots, Z_2, Z_1)'$,

Box and Jenkin (1976) suggested the unconditional log-likelihood function:

$$\ln L(\phi, \mu, \theta, \sigma_a^2) = -\frac{n}{2} \ln(2\pi\sigma_a^2) - \frac{S(\phi, \mu, \theta)}{2\sigma_a^2}$$

where $S(\phi, \mu, \theta) = \sum_{t=-\infty}^{n} [E(a_t|\phi, \mu, \theta, Z)]^2$, where $Z = (Z_1, Z_2, \cdots, Z_n)$ and $E(a_t|\phi, \mu, \theta, Z)$ is the conditional expectation of a_t given (ϕ, μ, θ, Z).

The maximizer of $\ln L(\phi, \mu, \theta, \sigma_a^2)$ is called the unconditional ML estimator (UMLE) of $(\phi, \mu, \theta, \sigma_a^2)$, denoted by $(\widehat{\phi}, \widehat{\mu}, \widehat{\theta}, \widehat{\sigma}_a^2)$.

$(\widehat{\phi}, \widehat{\mu}, \widehat{\theta})$ is also the minimizer of $S(\phi, \mu, \theta)$. So, $(\widehat{\phi}, \widehat{\mu}, \widehat{\theta})$ is called unconditional LS estimator (ULSE) of (ϕ, μ, θ).

$S(\phi, \mu, \theta)$ is approximated by

$$S(\phi, \mu, \theta) = \sum_{t=-M}^{n} [E(a_t|\phi, \mu, \theta, Z)]^2 \,,$$

where M is an integer large enough so that, for any predetermined $\epsilon > 0$,

$$|E(a_t|\phi, \mu, \theta, Z) - E(a_{t-1}|\phi, \mu, \theta, Z)| \leq \epsilon$$

for $t \leq -M - 1$.

How to find $E(a_t|\phi, \mu, \theta, Z)$? Backcasting method.

σ_a^2 is estimated by

$$\hat{\sigma}_a^2 = \frac{S(\hat{\phi}, \hat{\mu}, \hat{\theta})}{n}.$$

SAS program:

```
proc arima data=yenusa;

identify var=log(1) noprint;

estimate q=1 noconstant method=uls;

run;
```

Backcasting Method.

An ARMA model (forward form):

$(1-\phi_1 B-\cdots-\phi_p B^p)\dot{Z}_t=(1-\theta_1 B-\cdots-\theta_q B^q)a_t$
can be written by

An ARMA model (backward form):

$(1-\phi_1 F-\cdots-\phi_p F^p)\dot{Z}_t=(1-\theta_1 F-\cdots-\theta_q F^q)e_t\,,$
where $e_t \sim N(0,\sigma_e^2)$ and $F^j Z_t = Z_{t+j}$.

Example 7.4.

AR(1) model in foreward form: $(1-\phi_1 B)Z_t = a_t$.

AR(1) model in backward form: $(1-\phi_1 F)Z_t = e_t$.

Question: how to find $E(a_t|\phi, Z)$?

Solution:

$$
\begin{aligned}
&E(a_n|\phi,Z)=E(Z_n|Z)-\phi E(Z_{n-1}|Z)=Z_n-\phi Z_{n-}\\
&\vdots\\
&E(a_2|\phi,Z)=E(Z_2|Z)-\phi E(Z_1|Z)=Z_2-\phi Z_1\,,\\
&E(a_1|\phi,Z)=E(Z_1|Z)-\phi E(Z_0|Z)=Z_1-\phi E(Z_0|\\
&E(a_0|\phi,Z)=E(Z_0|Z)-\phi E(Z_{-1}|Z)\,,\\
&E(a_{-1}|\phi,Z)=E(Z_{-1}|Z)-\phi E(Z_{-2}|Z)\,,\\
&\vdots\\
&E(a_{-M}|\phi,Z)=E(Z_{-M}|Z)-\phi E(Z_{-M-1}|Z)
\end{aligned}
$$

Further Question: how to find $E(Z_{-t}|Z)$?

By backward form:

$$Z_t = \sum_{i=0}^{\infty} \phi^i F^i e_t = e_t + \phi e_{t+1} + \phi^2 e_{t+2} + \cdots \,,$$

$$\begin{aligned}
E(Z_0|Z) &= E(e_0|Z) + \phi E(Z_1|Z) = \phi Z_1 \,, \\
E(Z_{-1}|Z) &= E(e_{-1}|Z) + \phi E(Z_0|Z) = \phi^2 Z_1 \,, \\
&\vdots \\
E(Z_{-M}|Z) &= E(e_{-M}|Z) + \phi E(Z_{-M+1}|Z) = \phi^{M+1}
\end{aligned}$$

$E(Z_{-t}|Z)$ is called the backcasting of Z_{-t}.
Thus

$$\begin{aligned}
&E(a_1|\phi, Z) = Z_1 - \phi^2 Z_1 = (1-\phi^2) Z_1 \,, \\
&E(a_0|\phi, Z) = (1-\phi^2)\phi Z_1 \,, \\
&E(a_{-1}|\phi, Z) = (1-\phi^2)\phi^2 Z_1 \,, \\
&\vdots \\
&E(a_{-M}|\phi, Z) = (1-\phi^2)\phi^{M+1} Z_1.
\end{aligned}$$

Now,

$$S(\phi) = \sum_{t=2}^{n} \left[\dot{Z}_t - \phi Z_{t-1}\right]^2 + \sum_{i=-M}^{1} \left[(1-\phi^2)\phi^{1-i} Z_1\right]^2 .$$

7.2.3. Exact likelihood function

Model: $\dot{Z}_t = Z_t - \mu$,

$$\begin{aligned}\dot{Z}_t &= \phi_1 \dot{Z}_{t-1} + \phi_2 \dot{Z}_{t-1} + \cdots + \phi_p \dot{Z}_{t-p} \\ &\quad + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q} ,\end{aligned}$$

where $a_t \sim i.i.d.N(0, \sigma_a^2)$.

Given $Z_1, Z_2, \cdots, Z_n$,

the joint probability $f(Z_n, Z_{n-1}, \cdots, Z_1)$ is the exact likelihood function.

Let $L(\phi, \mu, \theta, \sigma_a^2) = f(Z_n, Z_{n-1}, \cdots, Z_1)$.

The maximizer of $\ln L(\phi, \mu, \theta, \sigma_a^2)$ is called the (exact) ML estimator (MLE) of $(\phi, \mu, \theta, \sigma_a^2)$, denoted by $(\widehat{\phi}, \widehat{\mu}, \widehat{\theta}, \widehat{\sigma}_a^2)$.

How to find the exact likelihood function?

Answer: too complicated.

Example 7.5. AR(1) model:

$$Z_t = \phi_1 Z_{t-1} + a_t \,,$$

where $|\phi| < 1$ and $a_t \sim N(0, \sigma_a^2)$.

Solution:

$$Z_t = \sum_{j=0}^{\infty} \phi^j a_{t-j}.$$

So, $Z_t \sim N(0, \sigma_a^2/(1-\phi^2))$. Consider

$$\begin{aligned} e_1 = Z_1 = \sum_{j=0}^{\infty} \phi^j a_{1-j} \,, \\ a_2, \\ a_3, \\ \vdots \\ a_n. \end{aligned}$$

Since $e_1, a_2, \cdots, a_n$ are independent, their joint density is

$$\begin{aligned} f(e_1, a_2, \cdots, a_n) \;&=\; \left[\frac{1-\phi^2}{2\pi\sigma_a^2}\right]^{1/2} e^{-\frac{e_1^2(1-\phi^2)}{2\,\sigma_a^2}} \\ &\cdot \left[\frac{1}{2\pi\sigma_a^2}\right]^{\frac{n-1}{2}} e^{-\frac{1}{2\,\sigma_a^2}\sum_{t=2}^{n} a_t^2} . \end{aligned}$$

Using the transformation:

$$
\begin{aligned}
&Z_1 = e_1, \\
&Z_2 = \phi Z_1 + a_2, \\
&Z_3 = \phi Z_2 + a_3, \\
&\vdots \\
&Z_n = \phi Z_{n-1} + a_n.
\end{aligned}
$$

The Jacobian for the transformation is

$$
J = \begin{vmatrix}
1 & 0 & \cdot & \cdot & \cdot & 0 & 0 & 0 \\
-\phi & 1 & 0 & \cdot & \cdot & 0 & 0 & 0 \\
0 & -\phi & 1 & 0 & \cdot & 0 & 0 & 0 \\
\vdots & & & & & & & \\
0 & \cdot & \cdot & \cdot & \cdot & 0 & -\phi & 1
\end{vmatrix} = 1.
$$

Thus

$$
\begin{aligned}
f(Z_n, Z_{n-1}, \cdots, Z_1) &= J f(e_1, a_2, \cdots, a_n) \\
&= \left[\frac{1-\phi^2}{2\pi\sigma_a^2}\right]^{\frac{1}{2}} e^{-\frac{Z_1^2(1-\phi^2)}{2\,\sigma_a^2}} \\
&\cdot \left[\frac{1}{2\pi\sigma_a^2}\right]^{\frac{n-1}{2}} e^{-\frac{1}{2\,\sigma_a^2}\sum_{t=2}^{n}(Z_t-\phi Z_t
\end{aligned}
$$

$$
\begin{aligned}
&\ln L(\phi, \sigma_a^2) \\
&= -\frac{n}{2}\ln(2\pi) + \frac{1}{2}\ln(1-\phi^2) - \frac{n}{2}\ln\sigma_a^2 - \frac{S(\phi)}{2\,\sigma_a^2}
\end{aligned}
$$

where

$$
S(\phi) = Z_1^2(1-\phi^2) + \sum_{t=2}^{n}(Z_t - \phi Z_{t-1})^2.
$$

Section 7.3 Nonlinear Estimation

(1) In CLS, we need to find $(\widehat{\phi}, \widehat{\mu}, \widehat{\theta})$ so that

$$S_*(\phi, \mu, \theta) = \sum_{t=1}^{n} a_t^2(\phi, \mu, \theta)$$

achieves its minimum value.

(2) In ULS, for example, for the AR(1) model, we need to find $\widehat{\phi}$ so that

$$S(\phi) = \sum_{t=2}^{n} \left[\dot{Z}_t - \phi Z_{t-1}\right]^2 + \sum_{i=-M}^{1} \left[(1-\phi^2)\phi^{1-i} Z_1\right]^2.$$

achieves its minimum value.

(3) In ML, for example, for the AR(1) model, we need to find $\widehat{\phi}$ so that

$$S(\phi) = Z_1^2(1-\phi^2) + \sum_{t=2}^{n} (Z_t - \phi Z_{t-1})^2.$$

achieves its minimum value.

Question: how to find?

nonlinear estimation method.

Example 7.1 ARMA(1, 1) model

$$Z_t = \phi Z_{t-1} + a_t - \theta a_{t-1}.$$

Find $(\widehat{\phi}, \widehat{\theta})$ so that

$$S_*(\phi, \theta) = \sum_{t=1}^{n} a_t^2(\phi, \theta)$$

achieves its minimum value.

Properties of the Parameter Estimates:

Let $\widehat{\alpha} = (\widehat{\phi}, \widehat{\mu}, \widehat{\theta})$ be the CLS, ULS, or ML estimator of $\alpha = (\phi, \mu, \theta)$. Then

$$\sqrt{n}(\widehat{\alpha} - \alpha) \sim N(0, V(\widehat{\alpha}))\,,$$

where

$$V(\widehat{\alpha}) = \widehat{\sigma}_a^2 \left(\bar{X}'_{\widehat{\alpha}} \bar{X}_{\widehat{\alpha}}\right)^{-1} = \left(\widehat{\sigma}^2_{\widehat{\alpha}_i \widehat{\alpha}_j}\right).$$

We can test the hypothesis $H_0 : \alpha_i = \alpha_{i0}$ using the following t statistics:

$$t = \frac{\widehat{\alpha}_i - \alpha_{i0}}{\widehat{\sigma}_{\widehat{\alpha}_i \widehat{\alpha}_i}}$$

with the degrees of freedom equaling $n - (p + q + 1)$ for the ARMA(p, q) model.

The estimated correlation matrix of these estimates is

$$\widehat{R}(\alpha) = \left(\widehat{\rho}_{\widehat{\alpha}_i \widehat{\alpha}_j}\right).$$

where

$$\widehat{\rho}_{\widehat{\alpha}_i \widehat{\alpha}_j} = \frac{\widehat{\sigma}_{\widehat{\alpha}_i \widehat{\alpha}_j}}{\sqrt{\widehat{\sigma}_{\widehat{\alpha}_i \widehat{\alpha}_i}}\sqrt{\widehat{\sigma}_{\widehat{\alpha}_j \widehat{\alpha}_j}}}$$

A high correlation among estimates indicates overparameterization (too many parameters).

Section 7.4. Omitted

Section 7.5. Diagnostic Checking

Given data: $Z_1, \cdots, Z_n$. Assume the model is ARMA(p, q):

$$\begin{aligned} Z_t &= \mu + \phi_1 Z_{t-1} + \cdots + \phi_p Z_{t-p} \\ &\quad + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}. \end{aligned}$$

Two problems need to be considered:

(i). Is the **ARMA** model correct?

(ii). Even if the **ARMA** model is correct, what is p and q?

Suppose (i) and (ii) are correct. We can estimated $\widehat{\phi}_i$ and $\widehat{\theta}_j$.

Let

$$\begin{aligned} \widehat{a}_t &= Z_t - \widehat{\mu} - \widehat{\phi}_1 Z_{t-1} - \cdots - \widehat{\phi}_p Z_{t-p} \\ &\quad + \widehat{\theta}_1 \widehat{a}_{t-1} + \cdots + \widehat{\theta}_q \widehat{a}_{t-q}, \end{aligned}$$

where $t = 1, \cdots, n$.

$\widehat{a}_t$, $t = 1, \cdots, n$, are called residuals.

If (i) and (ii) are truly correct, then $\widehat{a}_t$ should be very close to a_t.

Since the ACF of a_t is zero, the ACF of $\widehat{a}_t$ should be very close to zero. That is , we need to check the null hypothesis:

$$\begin{aligned} H_0: &\quad \rho_1 = \rho_2 = \cdots = \rho_K = 0. \\ H_1: &\quad H_0 \quad \text{does not holds.} \end{aligned}$$

Test statistics:

$$Q = n(n+2)\sum_{k=1}^{K}(n-k)^{-1}\widehat{\rho}_k^2 \sim \chi^2(K-m)\,,$$

K is the lag of ACF specified by yourself and m is the number of the parameters estimated in the model,

$$\widehat{\rho}_k = \sum_{t=1}^{n-k}\widehat{a}_t\widehat{a}_{t+k}\Bigg/\sum_{t=1}^{n}\widehat{a}_t^2.$$

Section 7.6. Empirical Examples for Series W1-W7

Section 7.7. Model Selection Criteria

1. Akaike's AIC and BIC Criteria

Akaike (1973, 1974):

$$\begin{aligned} \text{AIC}(p,q) &= \ln \widehat{\sigma}_a^2 + 2(p+q)/n, \\ \text{AIC}(p,q) &= -2 \ln L_*(\widehat{\phi}, \widehat{\mu}, \widehat{\theta}, \widehat{\sigma}_a^2) + 2(p+q) \end{aligned}$$

The optimal order (p,q) is the one so that AIC(p,q) is minimum.——- called AIC [Akaike's information criterion].

$$\begin{aligned} \text{BIC}(p,q) &= \ln \widehat{\sigma}_a^2 + 2(p+q) \ln n/n, \\ \text{BIC}(p,q) &= -2 \ln L_*(\widehat{\phi}, \widehat{\mu}, \widehat{\theta}, \sigma_a^2) + (p+q) \ln n \end{aligned}$$

The optimal order (p,q) is the one so that BIC(p,q) is minimum.——- called BIC [Schwartz (1978) Bayesian Criterion](SBC).