Lecture 9

ARMA Models: Properties, Identification, and Estimation

Reading: Bowerman, O'Connell, and Koehler (2005): Chapter 9.2-9.4; Capter 10.1; Cryer and Chen (2008): Chapter 4.4-4.6; Chapter 6.1-6.3

MATH 4070: Regression and Time-Series Analysis

Whitney Huang
Clemson University

Notes

Agenda

1. Properties of ARMA Models: Stationarity, Causality, and Invertibility
2. Tentative Model Identification Using ACF and PACF
3. Parameter Estimation

Notes

ARMA(p, q) Processes

$\{\eta_t\}$ is an ARMA(p, q) process if it satisfies

$$\eta_t - \sum_{i=1}^{p} \phi_i \eta_{t-i} = Z_t + \sum_{j=1}^{q} \theta_j Z_{t-j},$$

where $\{Z_t\}$ is a $\mathrm{WN}(0, \sigma^2)$ process.

- Let $\phi(B) = 1 - \sum_{i=1}^{p} \phi_i B^i$ and $\theta(B) = 1 + \sum_{j=1}^{q} \theta_j B^j$. Then we can write it as

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

- An ARMA(p, q) process $\{\tilde{\eta}_t\}$ with mean μ can be written as

$$\phi(B)(\tilde{\eta}_t - \mu) = \theta(B)Z_t$$

Notes

A Stationary Solution to the ARMA Equation

A zero-mean ARMA process is stationary if it can be written as a linear process, i.e., $\eta_t = \psi(B)Z_t$, where $\psi(B) = \sum_{j=-\infty}^{\infty} \psi_j B^j$ for an absolutely summable sequence $\{\psi_j\}$

- This only happens if one can "divide" by $\phi(B)$, i.e., it is stationary only if the following makes sense:

$$(\phi(B))^{-1}\phi(B)\eta_t = (\phi(B))^{-1}\theta(B)Z_t$$
$$\Rightarrow \eta_t = \underbrace{\frac{\theta(B)}{\phi(B)}}_{=\psi(B)} Z_t$$

- Let's forget about B is the backshift operator and replace it with z. Now consider whether we can divide $\theta(z)$ by $\phi(z)$

Notes

Roots of the AR Characteristic Polynomial and Stationarity

- A root of the polynomial $f(z) = \sum_{j=0}^{p} a_j z^j$ is a value ξ such that $f(\xi) = 0 \Rightarrow$ it can be real-valued $\mathbb{R}$ or complex-valued $\mathbb{C}$
- For example, a root can take the form $\xi = a + b\,i$ for real number a and b. The modulus of a complex number $|\xi|$ is defined by

$$|\xi| = \sqrt{a^2 + b^2}$$

- For any ARMA(p,q) process, a stationary and unique solution exists if and only if

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0,$$

for all $|z| = 1 \Rightarrow$ None of the roots of the AR characteristic equation have a modulus of exactly 1

Note: Stationarity of the ARMA process has nothing to do with the MA polynomial!

Notes

AR(4) Example

Consider the following AR(4) process

$$\eta_t = 2.7607\eta_{t-1} - 3.8106\eta_{t-2} + 2.6535\eta_{t-3} - 0.9238\eta_{t-4} + Z_t,$$

the AR characteristic polynomial is

$$\phi(z) = 1 - 2.7607z + 3.8106z^2 - 2.6535z^3 + 0.9238z^4$$

- Hard to find the roots of $\phi(z)$ –we use the `polyroot` function in `R`:
- Use `Mod` in `R` to calculate the modulus of the roots
- **Conclusion:**

Notes

Causal ARMA Processes

An ARMA process is causal if there exists constants $\{\psi_j\}$ with $\sum_{j=0}^{\infty} |\psi_j| < 0$ and $\eta_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$, that is, we can write $\{\eta_t\}$ as an MA(∞) process depending only on the current and past values of $\{Z_t\}$

- Equivalently, an ARMA process is causal if and only if

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0,$$

  for all $|z| \leq 1 \Rightarrow$ None of the roots of the AR characteristic equation have a modulus less than 1

- The previous AR(4) example is causal since each zero, ξ, of $\phi(\cdot)$ is such that $|\xi| > 1$

Note: The causality of the ARMA process depends only on the AR polynomial!

Properties of ARMA Models: Stationarity, Causality, and Invertibility

Tentative Model Identification Using ACF and PACF

Parameter Estimation

Notes

Invertible ARMA Processes

An ARMA process is invertible if there exists constants $\{\pi_j\}$ with $\sum_{j=0}^{\infty} |\pi_j| < \infty$ and

$$Z_t = \sum_{j=0}^{\infty} \pi_j \eta_{t-j},$$

that is, we can write $\{Z_t\}$ as an AR(∞) process depending only on the current and past values of $\{\eta_t\}$

- A process is invertible if and only if

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q \neq 0,$$

  for all $|z| \leq 1 \Rightarrow$ None of the roots of the MA characteristic equation have a modulus less than 1

- An ARMA process

$$\eta_t - 0.5\eta_{t-1} = Z_t + 0.4 Z_{t-1},$$

  with $\phi(z) = 1 - 0.5z$ and $\theta(z) = 1 + 0.4z$ has a root of the MA characteristic polynomial at $z = \frac{-1}{0.4} = -2.5$

Properties of ARMA Models: Stationarity, Causality, and Invertibility

Tentative Model Identification Using ACF and PACF

Parameter Estimation

Notes

Review of the Autocorrelation Function (ACF)

The autocorrelation function (ACF) measures the correlation of a stationary time series η_t with its own lagged values

- The theoretical ACF for MA processes can be computed as $\rho(h) = \frac{\sum_{j=0}^{q} \theta_j \theta_{j+h}}{\sum_{j=0}^{q} \theta_j^2}$, and via the Yule-Walker equation for AR processes

- The ACF is useful in identifying the MA(q) order, as it cuts off after lag q

Properties of ARMA Models: Stationarity, Causality, and Invertibility

Tentative Model Identification Using ACF and PACF

Parameter Estimation

Notes

Partial Autocorrelation Functions (PACF)

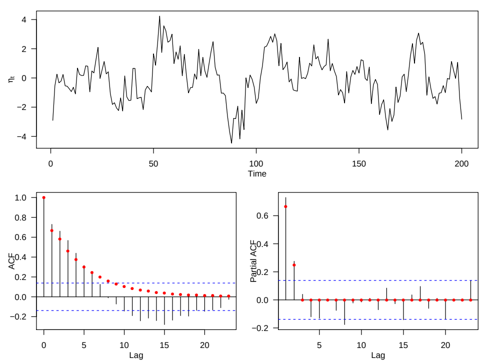
The partial autocorrelation function (PACF) represents the partial correlation of a stationary time series $\{\eta_t\}$ with its own lagged values, while regressing out the effects of the time series at all shorter lags

- The PACF at lag h is the autocorrelation between η_t and η_{t+h} with the linear dependence between η_t and $\eta_{t+1}, \ldots, \eta_{t+h-1}$ removed
- PACF plots are a commonly used tool for identifying the order of an AR model, as the theoretical PACF "shuts off" past the order of the model (see an example on the next slide)
- One can use the function `pacf` in `R` to plot the PACF

Notes

An Example of PACF Plot

$$\eta_t - 0.5\eta_{t-1} - 0.25\eta_{t-2} = Z_t$$

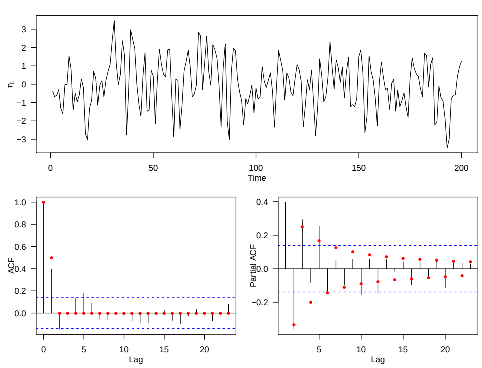


The theoretical ACF decays exponentially, while the PACF cuts off at lag 2

Notes

PACF Plot for a MA Process

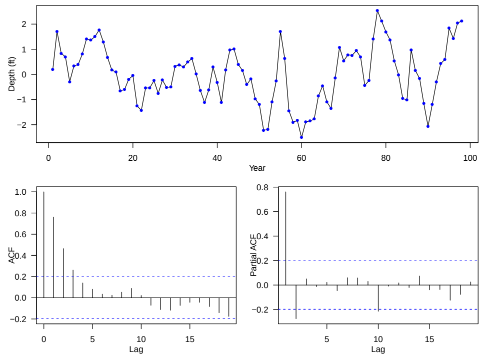
$$\eta_t = Z_t + Z_{t-1}$$



The theoretical ACF cuts off at lag 1, while the PACF decays exponentially

Notes

Lake Huron Series PACF Plot

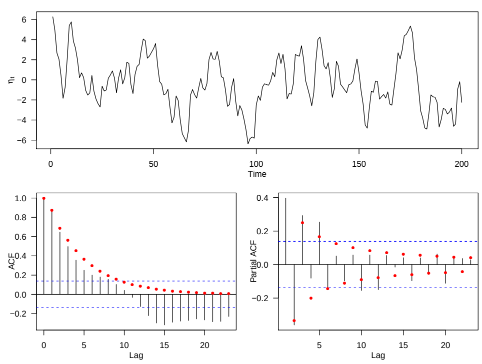


> We can use both ACF and PACF plots to identify the potential ARMA model order

Notes

PACF Plot for a ARMA Process

$$\eta_t - 0.5\eta_{t-1} - 0.25\eta_{t-2} = Z_t + Z_{t-1}$$



Both the theoretical ACF and PACF decay exponentially

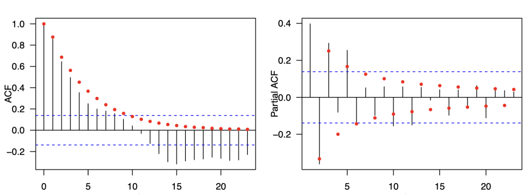
Notes

Identifying Plausible Stationary ARMA Models

We can use the sample ACF and PACF to help identify plausible models:

Model	ACF	PACF
MA(q)	cuts off after lag q	tails off exponentially
AR(p)	tails off exponentially	cuts off after lag p

For ARMA(p, q) we will see a combination of the above



Notes

Estimation of the ARMA Process Parameters

Suppose we choose a ARMA(p, q) model for $\{\eta_t\}$

- Need to estimate the $p+q+1$ parameters:
  - AR component $\{\phi_1, \cdots, \phi_p\}$
  - MA component $\{\theta_1, \cdots, \theta_q\}$
  - $\mathrm{Var}(Z_t) = \sigma^2$
- One strategy:
  - Do some preliminary estimation of the model parameters (e.g., via Yule-Walker estimates)
  - Follow-up with maximum likelihood estimation with Gaussian assumption

Notes

The Yule-Walker Method

Suppose η_t is a causal AR(p) process

$$\eta_t - \phi_1 \eta_{t-1} - \cdots - \phi_p \eta_{t-p} = Z_t$$

To estimate the parameters $\{\phi_1, \cdots, \phi_p\}$, we use a method of moments estimation scheme:

- Let $h = 0, 1, \cdots, p$. We multiply η_{t-h} to both sides

$$\eta_t \eta_{t-h} - \phi_1 \eta_{t-1} \eta_{t-h} - \cdots - \phi_p \eta_{t-p} \eta_{t-h} = Z_t \eta_{t-h}$$

- Taking expectations:

$$\mathbb{E}(\eta_t \eta_{t-h}) - \phi_1 \mathbb{E}(\eta_{t-1}\eta_{t-h}) - \cdots - \phi_p \mathbb{E}(\eta_{t-p}\eta_{t-h}) = \mathbb{E}(Z_t \eta_{t-h}),$$

  we get

$$\boxed{\gamma(h) - \phi_1 \gamma(h-1) - \cdots - \phi_p \gamma(h-p) = \mathbb{E}(Z_t \eta_{t-h})}$$

ARMA Models: Properties, Identification, and Estimation

School of MATHEMATICAL AND STATISTICAL SCIENCES Clemson University

Properties of ARMA Models: Stationarity, Causality, and Invertibility

Tentative Model Identification Using ACF and PACF

Parameter Estimation

Notes

The Yule-Walker Equations

- When $h = 0$, $\mathbb{E}(Z_t \eta_{t-h}) = \mathrm{Cov}(Z_t, \eta_t) = \sigma^2$ (Why?) Therefore, we have

$$\gamma(0) - \sum_{j=1}^{p} \phi_j \gamma(j) = \sigma^2$$

- When $h > 0$, Z_t is uncorrelated with η_{t-h} (because the assumption of causality), thus $\mathbb{E}(Z_t \eta_{t-h}) = 0$ and we have

$$\gamma(h) - \sum_{j=1}^{p} \phi_j \gamma(h-j) = 0, \quad h = 1, 2, \cdots, p$$

- The Yule-Walker estimates are the solution of these equations when we replace $\gamma(h)$ by $\hat{\gamma}(h)$

ARMA Models: Properties, Identification, and Estimation

School of MATHEMATICAL AND STATISTICAL SCIENCES Clemson University

Properties of ARMA Models: Stationarity, Causality, and Invertibility

Tentative Model Identification Using ACF and PACF

Parameter Estimation

Notes

The Yule-Walker Equations in Matrix Form

Let $\hat{\phi} = (\hat{\phi}_1, \cdots, \hat{\phi}_p)^T$ be an estimate for $\phi = (\phi_1, \cdots, \phi_p)^T$ and let

$$\hat{\mathbf{\Gamma}} = \begin{bmatrix} \hat{\gamma}(0) & \hat{\gamma}(1) & \cdots & \hat{\gamma}(p-1) \\ \hat{\gamma}(1) & \hat{\gamma}(0) & \cdots & \hat{\gamma}(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\gamma}(p-1) & \hat{\gamma}(p-2) & \cdots & \hat{\gamma}(0) \end{bmatrix}.$$

Then the Yule-Walker estimates of ϕ and σ^2 are

$$\hat{\phi} = \hat{\mathbf{\Gamma}}^{-1}\hat{\gamma},$$

and

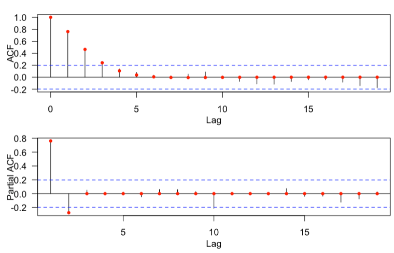
$$\hat{\sigma}^2 = \hat{\gamma}(0) - \hat{\phi}^T\hat{\gamma},$$

where $\hat{\gamma} = (\hat{\gamma}(1), \cdots, \hat{\gamma}(p))^T$

Notes

Lake Huron Example in R


```
```{r}
YW_est <- ar(lm$residuals, aic = F, order.max = 2, method = "yw")
# plot sample and estimated acf/pacf
par(las = 1, mgp = c(2.2, 1, 0), mar = c(3.6, 3.6, 0.6, 0.6), mfrow = c(2, 1))
acf(lm$residuals)
acf_YWest <- ARMAacf(ar = YW_est$ar, lag.max = 23)
points(0:23, acf_YWest, col = "red", pch = 16, cex = 0.8)
pacf(lm$residuals)
pacf_YWest <- ARMAacf(ar = YW_est$ar, lag.max = 23, pacf = T)
points(1:23, pacf_YWest, col = "red", pch = 16, cex = 0.8)
```
```



Notes

Remarks on the Yule-Walker Method

- For large sample size, Yule-Walker estimator have (approximately) the same sampling distribution as maximum likelihood estimator (MLE), but with small sample size Yule-Walker estimator can be far less efficient than the MLE
- The Yule-Walker method is a poor procedure for MA(q) and ARMA(p,q) processes with $q > 0$ (see Cryer Chan 2008, p. 150-151)
- We move on the more versatile and popular method for estimating ARMA(p,q) parameters–maximum likelihood estimation[1]

[1] See **Least Squares Estimation** in Chapter 7.2 of Cryer and Chan (2008).

Notes

Maximum Likelihood Estimation

- The setup:
  - Model: $\boldsymbol{X} = (X_1, X_2, \cdots, X_n)$ has joint probability density function $f(\boldsymbol{x}|\boldsymbol{\omega})$ where $\boldsymbol{\omega} = (\omega_1, \omega_2, \cdots, \omega_p)$ is a vector of p parameters
  - Data: $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)$
- The likelihood function is defined as the the "likelihood" of the data, $\boldsymbol{x}$, given the parameters, $\boldsymbol{\omega}$

$$L_n(\boldsymbol{\omega}) = f(\boldsymbol{x}|\boldsymbol{\omega})$$

- The maximum likelihood estimate (MLE) is the value of $\boldsymbol{\omega}$ which maximizes the likelihood, $L_n(\boldsymbol{\omega})$, of the data $\boldsymbol{x}$:

$$\hat{\boldsymbol{\omega}} = \underset{\boldsymbol{\omega}}{\operatorname{argmax}} L_n(\boldsymbol{\omega}).$$

  It is equivalent (and often easier) to maximize the log likelihood,

$$\ell_n(\boldsymbol{\omega}) = \log L_n(\boldsymbol{\omega})$$

Notes

The MLE for an i.i.d. Gaussian Process

Suppose $\{X_t\}$ be a Gaussian i.i.d. process with mean μ and variance σ^2. We observe a time series $\boldsymbol{x} = (x_1, \cdots, x_n)^T$.

- The likelihood function is

$$\begin{aligned} L_n(\mu, \sigma^2) &= f(\boldsymbol{x}|\mu, \sigma^2) \\ &= \prod_{t=1}^{n} f(x_t|\mu, \sigma) \\ &= \prod_{t=1}^{n} \left\{ \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x_t-\mu)^2}{2\sigma^2} \right] \right\} \\ &= (2\pi)^{-n/2}(\sigma^2)^{-n/2} \exp\left[-\frac{\sum_{t=1}^{n}(x_t-\mu)^2}{2\sigma^2} \right] \end{aligned}$$

- The log-likelihood function is

$$\begin{aligned} \ell_n(\mu, \sigma^2) &= \log L_n(\mu, \sigma^2) \\ &= -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{\sum_{t=1}^{n}(x_t-\mu)^2}{2\sigma^2} \end{aligned}$$

$$\Rightarrow \hat{\mu}_{\text{MLE}} = \frac{\sum_{t=1}^{n} X_t}{n} = \bar{X}, \quad \hat{\sigma}^2_{\text{MLE}} = \frac{\sum_{t=1}^{n}(X_t-\bar{X})^2}{n}$$

Notes

Likelihood for Stationary Gaussian Time Series Models

Suppose $\{X_t\}$ be a mean zero stationary Gaussian time series with ACVF $\gamma(h)$. If $\gamma(h)$ depends on p parameters, $\boldsymbol{\omega} = (\omega_1, \cdots, \omega_p)$

- The likelihood of the data $\boldsymbol{x} = (x_1, \cdots, x_n)$ given the parameters $\boldsymbol{\omega}$ is

$$L_n(\boldsymbol{\omega}) = (2\pi)^{-n/2}|\boldsymbol{\Gamma}|^{-1/2} \exp\left(-\frac{1}{2}\boldsymbol{x}^T\boldsymbol{\Gamma}^{-1}\boldsymbol{x} \right),$$

  where $\boldsymbol{\Gamma}$ is the covariance matrix of $\boldsymbol{X} = (X_1, \cdots, X_n)^T$, $|\boldsymbol{\Gamma}|$ is the determinant of the matrix $\boldsymbol{\Gamma}$, and $\boldsymbol{\Gamma}^{-1}$ is the inverse of the matrix $\boldsymbol{\Gamma}$

- The log-likelihood is

$$\ell_n(\boldsymbol{\theta}) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log|\boldsymbol{\Gamma}| - \frac{1}{2}\boldsymbol{x}^T\boldsymbol{\Gamma}^{-1}\boldsymbol{x}$$

Notes

Decomposing Joint Density into Conditional Densities

A joint distribution can be represented as the product of conditionals and a marginal distribution

- The simple version for $n = 2$ is:

$$f(x_1, x_2) = f(x_2|x_1)f(x_1)$$

- Extending for general n we get the following expression for the likelihood:

$$L_n(\boldsymbol{\theta}) = f(\boldsymbol{x};\boldsymbol{\theta}) = f(x_1)\prod_{t=2}^{n} f(x_t|x_{t-1},\cdots,x_1;\boldsymbol{\theta}),$$

and the log-likelihood is

$$\ell_n(\boldsymbol{\theta}) = \log f(\boldsymbol{x};\boldsymbol{\theta}) = \log(f(x_1))+\sum_{t=2}^{n}\log f(x_t|x_{t-1},\cdots,x_1;\boldsymbol{\theta}).$$

Notes

AR(1) Log-likelihood

Let $\{\eta_1, \eta_2, \cdots, \eta_n\}$ be a realization of a zero-mean stationary AR(1) Gaussian time series. Let $\boldsymbol{\theta} = (\phi, \sigma^2)$

$$\ell_n(\boldsymbol{\theta}) = \underbrace{\log(f(\eta_1))}_{\ell_{n,1}} + \underbrace{\sum_{t=2}^{n}\log f(\eta_t|\eta_{t-1},\cdots,\eta_1;\boldsymbol{\theta})}_{\ell_{n,2}}.$$

Note that for $t \geq 2$, $f(\eta_t|\eta_{t-1},\cdots,\eta_1) = f(\eta_t|\eta_{t-1})$, where $[\eta_t|\eta_{t-1}] \sim \mathrm{N}(\phi\eta_{t-1}, \sigma^2) \Rightarrow \ell_{n,2} =$

$$-\frac{(n-1)}{2}\log 2\pi - \frac{(n-1)}{2}\log\sigma^2 - \frac{\sum_{t=2}^{n}(\eta_t - \phi\eta_{t-1})^2}{2\sigma^2}$$

Also, we know $[\eta_1] \sim \mathrm{N}\left(0, \frac{\sigma^2}{(1-\phi^2)}\right) \Rightarrow \ell_{1,n} =$

$$\frac{-\log 2\pi}{2} - \frac{\log\sigma^2}{2} + \frac{\log(1-\phi^2)}{2} - \frac{(1-\phi^2)\eta_1^2}{2\sigma^2}$$

$$\Rightarrow \ell_n(\boldsymbol{\theta}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 - \frac{\sum_{t=2}^{n}(\eta_t - \phi\eta_{t-1})^2}{2\sigma^2} + \frac{\log(1-\phi^2)}{2} - \frac{(1-\phi^2)\eta_1^2}{2\sigma^2}$$

Notes

AR(1) Log-likelihood Cont'd

$$\ell_n(\boldsymbol{\theta}) = -\frac{n}{2}\log 2\pi - \frac{n}{2}\log\sigma^2 + \frac{\log(1-\phi^2)}{2} - \frac{S(\phi)}{2\sigma^2},$$

where $S(\phi) = \sum_{t=2}^{n}(\eta_t - \phi\eta_{t-1})^2 + (1-\phi^2)\eta_1^2$

- For given value of ϕ, $\ell_n(\phi, \sigma^2)$ can be maximized analytically with respect to σ^2

$$\hat{\sigma}^2 = \frac{S(\hat{\phi})}{n}$$

- Estimation of ϕ can be simplified by maximizing the **conditional sum-of-squares** $(\sum_{t=2}^{n}(\eta_t - \phi\eta_{t-1})^2)$

Notes

Likelihood Calculation via Innovations

Let the best linear one-step predictor of X_t be

$$\hat{X}_t = \begin{cases} 0, & t = 1; \\ P_{t-1}X_t, & t = 2, \cdots, n \end{cases}$$

- The one-step prediction errors or innovations are defined
$$U_t = X_t - \hat{X}_t, \quad t = 1, \cdots, n,$$
and the associated mean squared error is
$$\nu_{t-1} = \mathbb{E}\left[(X_t - \hat{X}_t)^2\right] = \mathbb{E}(U_t^2), \quad t = 1, \cdots, n.$$
- For a causal ARMA process we can write $\nu_{t-1} = \sigma^2 r_{t-1}$, where r_t only depends on the AR and MA parameters $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$, but not σ^2

Notes

Working with the Innovations

- Result I: $\{U_t\}$ is an independent set of RVs with
$$U_t \sim \mathrm{N}(0, \nu_{t-1}), t = 1, \cdots, n$$
$\Rightarrow$ the one-step prediction errors are uncorrelated with one another, and each each a normal distribution
- Result II: The likelihoods are the same if we use a model based on realizations of $\{X_t\}$ or a model based on realizations of $\{U_t\}$
- Therefore
$$\ell_n(\boldsymbol{\omega}) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\sum_{t=1}^{n}\log(\nu_{t-1}) - \frac{1}{2}\sum_{t=1}^{n}\left(\frac{u_t^2}{\nu_{t-1}}\right).$$
For a causal ARMA process this becomes
$$\ell_n(\boldsymbol{\phi}, \boldsymbol{\theta}, \sigma^2) = -\frac{n}{2}\log(2\pi) - \frac{n}{2}\log(\sigma^2) - \frac{1}{2}\sum_{t=1}^{n}\log(r_{t-1}) - \frac{1}{2\sigma^2}\sum_{t=1}^{n}\left(\frac{u_t^2}{r_{t-1}}\right)$$

ARMA Models: Properties, Identification, and Estimation
School of Mathematical and Statistical Sciences Clemson University
Properties of ARMA Models: Stationarity, Causality, and Invertibility
Tentative Model Identification Using ACF and PACF
Parameter Estimation

Notes

The MLEs of σ^2, ϕ, and θ

- Now take the derivative of ℓ_n with respect to σ^2, setting the derivative equal to zero and solving for $\sigma^2 \Rightarrow$
$$\hat{\sigma}^2 = \frac{S(\boldsymbol{\phi}, \boldsymbol{\theta})}{n},$$
where
$$S(\boldsymbol{\phi}, \boldsymbol{\theta}) = \sum_{t=1}^{n}\left(\frac{u_t^2}{r_{t-1}}\right).$$
- Substituting $\hat{\sigma}^2$ into ℓ_n, the MLE estimates of $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$, denoted by $\hat{\boldsymbol{\phi}}$ and $\hat{\boldsymbol{\theta}}$, respectively, are those values which **maximize**
$$\tilde{\ell}_n(\boldsymbol{\phi}, \boldsymbol{\theta}, \hat{\sigma}^2) = -\frac{n}{2}\log(\hat{\sigma}^2) - \frac{1}{2}\sum_{t=1}^{n}\log(r_{t-1}) - \frac{1}{2\hat{\sigma}^2}\sum_{t=1}^{n}\left(\frac{u_t^2}{r_{t-1}}\right)$$

ARMA Models: Properties, Identification, and Estimation
School of Mathematical and Statistical Sciences Clemson University
Properties of ARMA Models: Stationarity, Causality, and Invertibility
Tentative Model Identification Using ACF and PACF
Parameter Estimation

Notes