STAT302: Time Series Analysis

Chapter 4. Stationary Processes and ARMA Models

Sangbum Choi, Ph.D

Department of Statistics, Korea University

Outline

Introduction

- Develop classes of linear models for stationary time series.
- Study the properities of these models, especially features of autocorrelation function (ACF) and partial autocorrelation (PACF).
- Three classes of models are:
  - Moving average (MA) models
  - Autoregressive (AR) models
  - Autoregressive moving average (ARMA) models

Finite-order MA models

A process $\{X_t\}$ is moving average of order q, denoted as MA(q):

$$\begin{aligned} X_t &= \mu + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \cdots - \theta_q \epsilon_{t-q} \\ &= \mu + \epsilon_t - \sum_{i=q}^{q} \theta_i \epsilon_{t-i} \\ &= \mu + \left(1 - \sum_{i=q}^{q} \theta_i B^i \right) \epsilon_t = \mu + \theta\left(\mathbf{B}\right) \epsilon_t, \end{aligned}$$

where $\{\epsilon_t\}$ is a white noise process (i.e. ϵ_t are independent RVs with mean 0 and variance σ^2 for all t), $\theta_1, \theta_2, ..., \theta_q$ are MA coefficients.

Basic properties of finite-order MA models

- Let $\{X_t\}$ denote an MA(q) process.
- The process is obviously stationary.
- The autocovariance function (ACVF) is

$$\gamma(0) = \mathrm{Var}(X_t) = \sigma^2(1 + \theta_1^2 + \theta_2^2 + \cdots + \sigma_q^2)$$

for $k = 1, 2, ..., q$, and

$$\gamma(k) = 0 \quad \mathit{for}\ k > q.$$

Basic properties of finite-order MA models

- Thus the ACF is

$$\rho(k) = \text{Cor}(X_t, X_{t+k}) = \frac{\gamma(k)}{\gamma(0)}$$
$$= \begin{cases} 1 & \text{if} \quad k = 0 \\ \frac{-\theta_k + \theta_1\theta_{k+1} + \cdots + \theta_{q-k}\theta_q}{1 + \theta_1^2 + \cdots + \theta_q^2} & \text{if} \quad k = 1, 2, ..., q \\ 0 & \text{if} \quad k > q \end{cases}$$

- Note that $\rho(k) = 0$ for $k > q$ is a distinctive feature of the ACF of an MA(q) process.
- We say that the ACF "cuts off" after lag q (i.e. becomes 0).

MA(1) process

- With $q = 1$, we have the MA(1) equation

$$X_t = \mu + \epsilon_t - \theta\epsilon_{t-1} = \mu + (1 - \theta B)\epsilon_t$$

- The autocovariance function is

$$\gamma(k) = \begin{cases} \sigma^2(1+\theta^2) & \text{if} \quad k = 0 \\ \sigma^2(-\theta) & \text{if} \quad k = 1 \\ 0 & \text{if} \quad k > 1 \end{cases}$$

- The ACF is

$$\rho(k) = \begin{cases} 1 & \text{if} \quad k = 0 \\ -\frac{\theta}{1+\theta^2} & \text{if} \quad k = 1 \\ 0 & \text{if} \quad k > 1 \end{cases}$$

- The ACF cuts off after lag 1.

MA(1) process

- Suppose $\mu = 0, \theta = 0.8$ in an MA(1) process

$$X_t = \epsilon_t - 0.8\epsilon_{t-1}$$

- Then we have

$$\rho(1) = \frac{-0.8}{1 + 0.8^2} \approx -0.49$$

- Note that, for any θ,

$$|\rho(1)| = \frac{|\theta|}{1 + \theta^2} \leq \frac{1}{2}$$

and thus the restriction is beyond the usual $|\rho(1)| \leq 1$.

MA(1) process and invertibility

- Conversely, given an MA(1) process with $\rho(1) = -0.49$, what is the value θ?
- Solving the following equation for θ

$$-0.49 = \frac{-\theta}{1+\theta^2} \quad \textit{or} \quad (-0.49)\theta^2 + \theta + (-0.49) = 0$$

we have

$$\theta = \frac{-1 + \sqrt{1 - 4(-0.49)^2}}{2(-0.49)} \approx 0.8$$

- There are two solutions. What are they?

MA(1) process and invertibility

- In general, solving the equation for θ

$$\rho(1)\theta^2 + \theta + \rho(1) = 0$$

  we have

$$\theta = \frac{-1 \pm \sqrt{1 - 4\rho(1)^2}}{2\rho(1)}$$

- There are two solutions if $|\rho(1)| \leq 1/2$ (considered as admissible values of $\rho(1)$).
- Denote the two solutions as θ, θ^*. Then $\theta = 1/\theta^*$ where $|\theta| < 1$ and $|\theta^*| > 1$.

MA(1) process and invertibility

- By "convention", to have a one-to-one correspondence between values of $\rho(1)$ and θ, always take θ such that $|\theta| < 1$.
- That is, there is usually two choices for MA(1) model:

$$X_t = \mu + \epsilon_t - \theta\epsilon_{t-1}$$

  with $|\theta| < 1$ and $|\theta| > 1$.
- It turns out that the MA model with $|\theta| < 1$ is invertible to the model with an infinite AR representation.

MA(2) process

- With $q = 2$, we have the MA(2) equation

$$\begin{aligned} X_t &= \mu + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} \\ &= \mu + (1 - \theta_1 B - \theta_2 B^2)\epsilon_t \\ &= \mu + \theta(B)\epsilon_t \end{aligned}$$

- The autocovariance function is

$$\gamma(k) = \begin{cases} \sigma^2(1 + \theta_1^2 + \theta_2^2) & \text{if} \quad k = 0 \\ \sigma^2(-\theta_1 + \theta_1\theta_2) & \text{if} \quad k = 1 \\ \sigma^2(-\theta_2) & \text{if} \quad k = 2 \\ 0 & \text{if} \quad k > 2 \end{cases}$$

MA(2) process

- The ACF is

$$\rho(k) = \begin{cases} 1 & \text{if} \quad k = 0 \\ \frac{-\theta_1 + \theta_1\theta_2}{1+\theta_1^2+\theta_2^2} & \text{if} \quad k = 1 \\ \frac{-\theta_2}{1+\theta_1^2+\theta_2^2} & \text{if} \quad k = 2 \\ 0 & \text{if} \quad k > 2 \end{cases}$$

- The ACF cuts off after lag 2.

Outline

Finite-order MA Models

Finite-order AR Models

Finite-order ARMA Models

Data Examples with R

Finite-order AR models

A process $\{X_t\}$ is autoregressive of order p, donoted as AR(p), if it satisfies

$$\begin{aligned} X_t &= \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \delta + \epsilon_t \\ &= \sum_{i=1}^{p} \phi_i X_{t-i} + \delta + \epsilon_t \end{aligned}$$

where $\{\epsilon_t\}$ is a white noise process, $\phi_1, \phi_2, ..., \phi_p$ are AR coefficients, and δ is a constant.

Finite-order AR models

It can be rewritten as

$$
\begin{aligned}
& X_t - \sum_{i=1}^{p} \phi_i X_{t-i} = \delta + \epsilon_t \\
\Longleftrightarrow \quad & \left(1 - \sum_{i=1}^{p} \phi_i B^i\right) X_t = \delta + \epsilon_t \\
\Longleftrightarrow \quad & \phi(B) X_t = \delta + \epsilon_t
\end{aligned}
$$

where

$$\phi(B) = 1 - \sum_{i=1}^{p} \phi_i B^i$$

is an AR(p) operator or often called a **characteristic function**.

Basic properties of finite-order AR models

- Let $\{X_t\}$ denote an AR(p) process. The process may or may not be stationary.
- Main issues to consider:
  1. Determnine conditions under which a process will be stationary.
  2. Under stationarity, develop characteristics of the autocovariance function and ACF.
- A key strategy is to attempt to represent $\{X_t\}$ by a linear process in MA(∞) form

$$X_t = \mu + \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$$

  where $\{\epsilon_t\}$ is a white noise process and $\sum_{j=0}^{\infty} |\psi_j| < \infty$.
- Then, $\{X_t\}$ is stationary.

AR(1) process

- With $p = 1$, we have the AR(1) equation

$$X_t = \phi X_{t-1} + \delta + \epsilon_t$$

- Now express $\{X_t\}$ in an MA(∞) form by successive substitutions of

$$X_{t-j} = \phi X_{t-j-1} + \delta + \epsilon_{t-j}, \quad j = 1, 2, ...,$$

into the AR(1) equation.

AR(1) process

- Thus we have, if $|\phi| < 1$,

$$X_t = \frac{\delta}{1-\phi} + \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} = \mu + \psi(B)\epsilon_t,$$

where $\psi_j = \phi^j$, $\psi(B) = \sum_{j=0}^{\infty} \phi^j B^j$, and

$$\sum_{j=0}^{\infty} |\psi_j| = \sum_{j=0}^{\infty} |\phi^i| = \frac{1}{1-|\phi|} < \infty.$$

- We conclude that if $|\phi| < 1$, then the process $\{X_t\}$ satisfying

$$X_t = \phi X_{t-1} + \delta + \epsilon_t$$

will be a stationary process. We call $|\phi| < 1$ the condition for stationarity of AR(1).

AR(1) process

- Assuming $|\phi| < 1$, we can derive the autocovariance function and ACF for AR(1).
- By the fact regarding a linear process, the autocovariance function is

$$\begin{aligned}\gamma(k) = \mathrm{Cov}(X_t, X_{t+k}) &= \sigma^2 \sum_{i=0}^{\infty} \psi_i \psi_{k+i} \\ &= \sigma^2 \left(\frac{\phi^k}{1-\phi^2} \right), \quad k = 0, 1, 2, \ldots\end{aligned}$$

- The ACF is

$$\rho(k) = \mathrm{Cor}(X_t, X_{t+k}) = \phi^k, k = 0, 1, 2, \ldots$$

AR(1) process

- Consider an alternative approach, which generalizes better to higher order AR processes.
- For $k \geq 0$, consider

$$\begin{aligned}\gamma(k) = \mathrm{Cov}(X_t, X_{t-k}) &= \mathrm{Cov}(\phi X_{t-1} + \delta + \epsilon_t, X_{t-k}) \\ &= \phi \mathrm{Cov}(X_{t-1}, X_{t-k}) + \mathrm{Cov}(\epsilon_t, X_{t-k}) \\ &= \phi\gamma(k-1) + \begin{cases} \sigma^2 & \text{if} \quad k = 0 \\ 0 & \text{if} \quad k > 0 \end{cases}\end{aligned}$$

- Thus we have

$$\gamma(k) = \phi\gamma(k-1), \;\; k = 1, 2, ...$$

AR(1) process

- Therefore,

$$\rho(k) = \phi\rho(k-1), k = 1, 2, ...$$

which are known as the Yule-Walker equations for AR(1) process.

- Thus the explicit solution of $\rho(k)$ is

$$\rho(k) = \phi\rho(k-1) = \phi^2\rho(k-2) = \cdots = \phi^{k-1}\rho(1) = \phi^k$$

- Also note that

$$\gamma(0) = \phi\gamma(1) + \sigma^2 = \phi^2\gamma(0) + \sigma^2$$
$$\Rightarrow \gamma(0) = \frac{\sigma^2}{1-\phi^2}$$

AR(p) process

- As with AR(1), express $\{X_t\}$ in an MA(∞) from by successive substitutions of $X_{t-1}, X_{t-2}, \ldots$ into the AR(p) equation

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \cdots + \phi_p X_{t-p} + \delta + \epsilon_t$$

- Starting from

$$\phi(B) X_t = \delta + \epsilon_t,$$

apply $\phi(B)^{-1}$ operator to both sides of the equation to obtain

$$X_t = \phi(B)^{-1}\delta + \phi(B)^{-1}\epsilon_t = \mu + \psi(B)\epsilon_t$$

where

$$\psi(B) = \phi(B)^{-1} = \sum_{j=0}^{\infty} \psi_j B^j$$

AR(p) process

- The coefficients $\{\psi_j\}$ are determined by the relation

$$\phi(B)\psi(B) = 1$$

- That is

$$\begin{aligned} & (1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p)(\psi_0 + \psi_1 B + \psi_2 B^2 + \cdots) \\ =& \psi_0 + (\psi_1 - \phi_1 \psi_0) B + (\psi_2 - \phi_1 \psi_1 - \phi_2 \psi_0) B^2 + \cdots \\ & \quad + (\psi_j - \phi_1 \psi_{j-1} - \phi_2 \psi_{j-2} - \cdots - \phi_p \psi_{j-p}) B^j + \cdots \\ =& \ 1 \end{aligned}$$

- It follows that

$$\psi_0 = 1, \ \psi_1 - \phi_1 \psi_0 = 0, \ \psi_2 - \phi_1 \psi_1 - \phi_2 \psi_0 = 0, \ \text{etc}$$

AR(p) process

- In general $\{\psi_j\}$ must satisfy, for $j = 1, 2, ...$,

$$\psi_j - \phi_1 \psi_{j-1} - \phi_2 \psi_{j-2} - \cdots - \phi_p \psi_{j-p} = 0$$

- Also

$$\psi_0 = 1, \ \psi_j = 0, \ j < 0$$

- These recursive relations implicitly determine the values of ψ_j.
- From the theory of solution to **difference equations**, we "know" that the explicit form of solutions for ψ_j depends on the p roots of the associated poly in variabel m:

$$m^p - \phi_1 m^{p-1} - \phi_2 m^{p-2} - \cdots - \phi_p = 0$$

AR(p) process

- If $|m_i| < 1$ for each $i = 1, 2, ..., p$, then the process $\{X_t\}$ is stationary and has infinite MA representation

$$X_t = \mu + \psi(B)\epsilon_t$$

where m_i are the roots of

$$m^p - \phi_1 m^{p-1} - \phi_2 m^{p-2} - \cdots - \phi_p = 0,$$
$$\psi(B) = \phi(B)^{-1} = \sum_{j=0}^{\infty} \psi_j B^j, \;\; \sum_{j=0}^{\infty} |\psi_j| < \infty$$

and $\{\psi_t\}$ satisfies

$$\psi_j - \phi_1\psi_{j-1} - \phi_2\psi_{j-2} - \cdots - \phi_p\psi_{j-p} = 0, \;\; j = 1, 2, ...$$
$$\psi_0 = 1, \;\; \psi_j = 0, j < 0$$

AR(p) process

- Condition for stationarity: All roots $m_1, m_2, ..., m_p$ of

$$m^p - \phi_1 m^{p-1} - \phi_2 m^{p-2} - \cdots - \phi_p = 0$$

  are less than 1 in absolute value.
- Equivalently, all roots $G_1, G_2, ..., G_p$ of

$$1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p = 0$$

  viewed as a poly in B must be greater than 1 in absolute value, since $G_i = 1/m_i$ for $i = 1, 2, ..., p$.

AR(2) process

- Suppose AR(2) with equation

$$X_t = X_{t-1} - 0.89X_{t-2} + \delta + \epsilon_t$$

  with $\phi_1 = 1.0$, $\phi_2 = -0.89$, and $\phi(B) = 1 - B + 0.89B^2$.
- The roots m_1, m_2 of $m^2 - m + 0.89 = 0$ are

$$\frac{1 \pm \sqrt{1 - 4(0.89)}}{2} = 0.5 \pm 0.8i$$

- Since

$$|m_1| = |m_2| = \sqrt{0.5^2 + 0.8^2} \approx 0.941 < 1$$

  the process $\{X_t\}$ is stationary.

AR(2) process

- The coefficients $\{\psi_j\}$ in the infinite MA representation of $\{X_t\}$ satisfy

$$\psi_j - \psi_{j-1} + 0.89\psi_{j-2} = 0, \;\; j = 1, 2, ...$$
$$\psi_0 = 1, \;\; \psi_j = 0, \;\; j < 0$$

- Since $\psi_j = \psi_{j-1} - 0.89\psi_{j-2}$ for $j = 1, 2, ...$, determine the values of ψ_j recursively as

$$\begin{aligned}
\psi_0 &= 1 \\
\psi_1 &= \psi_0 - 0.89\psi_{-1} = 1 \\
\psi_2 &= \psi_1 - 0.89\psi_0 = 1 - 0.89(1) = 0.11 \\
\psi_3 &= \psi_2 - 0.89\psi_1 = 0.11 - 0.89(1) = -0.78 \\
\psi_4 &= \psi_3 - 0.89\psi_2 = 0.78 - 0.89(0.11) = -0.878
\end{aligned}$$

Mean of AR(p)

- Let $\{X_t\}$ denote a stationary AR(p) process with the equation

$$X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + \delta + \epsilon_t$$

- Thus,

$$\mu = E(X_t) = \frac{\delta}{1 - \sum_{i=1}^{p} \phi_i}$$

- Note that if $\sum_{i=1}^{p} \phi_i = 1$, then m_i could be equal to 1, which violates the condition of stationarity.

Autocovariance function of AR(p)

- The autocovariance function (ACVF) can be obtained from the Yule-Walker equations.
- Note that, for $k \geq 0$,

$$\begin{aligned}
\gamma(k) &= \mathrm{Cov}(X_t, X_{t-k}) \\
&= \phi_1\gamma(k-1) + \phi_2\gamma(k-2) + \cdots + \phi_p\gamma(k-p) \\
&\qquad + \begin{cases} \sigma^2 & \text{if} \quad k = 0 \\ 0 & \text{if} \quad k > 0 \end{cases}
\end{aligned}$$

Autocovariance function of AR(p)

- Thus, for $k = 0$,

$$\gamma(0) = \mathrm{Var}(X_t) = \phi_1\gamma(1) + \phi_2\gamma(2) + \cdots + \phi_p\gamma(p) + \sigma^2$$

- Note that

$$\sigma^2 = \gamma(0)\left[1 - \phi_1\rho(1) - \phi_2\rho(2) - \cdots - \phi_p\rho(\rho)\right]$$

provides a relation between $\gamma(0) = \mathrm{Var}(X_t)$ and $\sigma^2 = \mathrm{Var}(\epsilon_t)$.

- Furthermore, for $k \geq 1$, the autocovariance function satisfies recursive relation

$$\gamma(k) = \phi_1\gamma(k-1) + \phi_2\gamma(k-2) + \cdots + \phi_p\gamma(k-p)$$

Yule-Walker equation

- Thus, for $k \geq 1$,

$$\rho(k) = \phi_1 \rho(k-1) + \phi_2 \rho(k-2) + \cdots + \phi_p \rho(k-p)$$

  which are known as the Yule-Walker (YW) equations for AR(p) processes.
- The YW equations can be used in two ways.
  1. To determine the ACF values $\rho(k), k = 1, 2, ...$ for given AR(p) coefficients $\phi_i, i = 1, 2, ..., p$.
  2. To determine the AR(p) coefficients $\phi_i, i = 1, 2, ..., p$ given the ACF values $\rho(k), k = 1, 2, ..., p$ from an AR(p) model.

$\phi \to \rho$: AR(2)

- Suppose AR(2) with equation

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \delta + \epsilon_t$$

- The YW equations are

$$\rho(k) = \phi_1 \rho(k-1) + \phi_2 \rho(k-2), \;\; k = 1, 2, \ldots$$

- Thus

$$\begin{aligned}
\rho(1) &= \phi_1 \rho(0) + \phi_2 \rho(1) \Rightarrow \rho(1) = \frac{\phi_1}{1-\phi_2} \\
\rho(2) &= \phi_1 \rho(1) + \phi_2 \rho(0) \\
\rho(3) &= \phi_1 \rho(2) + \phi_2 \rho(1) \\
&\cdots\cdots
\end{aligned}$$

- Recall AR(2) with $X_t = X_{t-1} - 0.89X_{t-2} + \delta + \epsilon_t$ with $\phi_1 = 1, \phi_2 = -0.89$.
- The roots of $m^2 - m + 0.89 = 0$ are $m_1, m_2 = 0.5 \pm 0.8i$, $|m_1| = |m_2| \approx 0.941 < 1$. Thus the process is stationary.
- The YW equations are

$$\rho(k) = \rho(k-1) - 0.89\rho(k-2), \quad k = 1, 2, ...$$

$\phi \to \rho$: AR(2)

- Thus

$$\rho(1) = \frac{\phi_1}{1-\phi_2} = \frac{1}{1-(-0.89)} = 0.529$$
$$\rho(2) = \phi_1\rho(1) + \phi_2\rho(0) = 0.529 - 0.89(1) = -0.361$$
$$\rho(3) = \phi_1\rho(2) + \phi_2\rho(1) = -0.361 - 0.89(0.529) = -0.832$$
$$\cdots\cdots$$

- Also we have

$$\gamma(0) = \frac{\sigma^2}{1-\phi_1\rho(1)-\phi_2\rho(2)} = \frac{\sigma^2}{1-0.529-(-0.89)(-0.361)}$$

$\phi \to \rho$: AR(p)

- For general AR(p), to determine $\rho(1), \rho(2), \ldots$ for given $\phi_1, \phi_2, \ldots \phi_p$, it is necessary to first solve a system of $p-1$ equations for $\rho(1), \ldots, \rho(p-1)$, then obtain $\rho(p), \rho(p+1), \ldots$ by simple recursion.
- For example with AR(3), first solve for $\rho(1), \rho(2)$,

$$\rho(1) = \phi_1 + \phi_2\rho(1) + \phi_3\rho(2)$$
$$\rho(2) = \phi_1\rho(1) + \phi_2 + \phi_3\rho(1)$$

- Then for $k \geq 3$,

$$\rho(k) = \phi_1\rho(k-1) + \phi_2\rho(k-2) + \phi_3\rho(k-3)$$

$\rho \to \phi$: AR(2)

- Now given $\rho(1), \rho(2)$, determine ϕ_1,ϕ_2 using the YW equations.
- Solving the following equations for ϕ_1, ϕ_2.

$$\rho(1) = \phi_1\rho(0) + \phi_2\rho(1)$$
$$\rho(2) = \phi_1\rho(1) + \phi_2\rho(0)$$

- Thus we have

$$\begin{cases} \phi_1 = \frac{\rho(1)[1-\rho(2)]}{1-\rho(1)^2} \\ \phi_2 = \frac{\rho(2)-\rho(1)^2}{1-\rho(1)^2} \end{cases}$$

$\rho \to \phi$: AR(p)

For a general p, the system of the first p YW equations is

$$\begin{bmatrix} 1 & \rho(1) & \rho(2) & \cdots & \rho(p-1) \\ \rho(1) & 1 & \rho(1) & \cdots & \rho(p-2) \\ \rho(2) & \rho(1) & 1 & \cdots & \rho(p-3) \\ \cdots & & & & \\ \rho(p-1) & \rho(p-2) & \rho(p-3) & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \cdots \\ \phi_p \end{bmatrix} = \begin{bmatrix} \rho(1) \\ \rho(2) \\ \rho(3) \\ \cdots \\ \rho(p) \end{bmatrix}$$

- Or

$$R_{p\times p}\phi_{p\times 1} = \rho_{p\times 1} \Rightarrow \phi = R^{-1}\rho$$

which can be used to determine ϕ for given $\rho(1), \ldots, \rho(p)$ in AR(p).

$\rho \to \phi$: AR(p)

- Recall the relation in AR(p)

$$\gamma(0) = \phi_1\gamma(1) + \phi_2\gamma(2) + \cdots + \phi_p\gamma(p) + \sigma^2$$

- Thus

$$R^2 = \frac{\phi_1\gamma(1) + \phi_2\gamma(2) + \cdots + \phi_p\gamma(p)}{\gamma(0)}$$

could be interpreted as the proportion of total variation that is explained by the AR(p) model (i.e. a measure of strength of dependency of X_t on past values $X_{t-1}, X_{t-2}, \cdots, X_{t-p}$).

- Furthermore,

$$\sigma^2 = \gamma(0)\left[1 - \phi_1\rho(1) - \phi_2\rho(2) - \cdots - \phi_p\rho(p)\right]$$

can be used to determine $\sigma^2 = \text{Var}(\epsilon_t)$ for given $\rho(1), ..., \rho(p)$, and $\rho(0)$.

Explicit form of AR(p) ACF

- Recall that the ACF of AR(p) satisfies the YW equations

  $$\rho(j) - \phi_1\rho(j-1) - \phi_2\rho(j-2) - \cdots - \phi_j\rho(j-p) = 0, \ j = 1, 2, \ldots$$

  which are in the form of a p^{th} order difference equation.
- In general, we have

  $$\rho(j) = c_1 m_1^j + c_2 m_2^j + \cdots + c_p m_p^j, \ \ j = 1, 2, \ldots$$

  where some roots are complex conjugates.

Partial autocorrelation function (PACF)

- PACF is a quantify defined with the motivation to identify the order p of an AR process, given the ACF of the process.
- In a practical situation, given observed data $X_1, X_1, \ldots, X_T$, if we consider fitting (or estimating) an AR model to the data, one issue would be to determine/choose/select the appropriate order p of the AR, and also to judge whether AR of any order is even appropriate.
- Let $\{X_t\}$ denote a stationary process with (known) ACF $\{\rho(k)\}$.
- If $\{X_t\}$ were an AR process of a particular order p, we want to define a quantify that is a "distinguishing" feature to characterize the order p of the AR process.

Partial autocorrelation function (PACF)

- For any value $k = 1, 2, 3, \ldots$ (which corresponds to possible order of the AR process), consider the $k \times k$ system of YW equations

$$\rho(j) = \phi_{1k}\rho(j-1) + \phi_{2k}\rho(j-2) + \cdots + \phi_{kk}\rho(j-k), \quad j = 1, 2, \ldots, k$$

- The partial autocorrelation coeffcient at lag k is ϕ_{kk}, which is the k^{th}(i.e. last) coefficient in the solution to the $k \times k$ system of YW equations, for $k = 1, 2, \ldots$
- The collection $\{\phi_{kk}\}_{k=1}^{\infty}$ is the partial autocorrelation function (PACF).

Partial autocorrelation function (PACF)

In matrix notation,

$$R_{k\times k}\phi_{k\times 1} = \rho_{k\times 1}$$

where

$$\begin{bmatrix} 1 & \rho(1) & \rho(2) & \cdots & \rho(k-1) \\ \rho(1) & 1 & \rho(1) & \cdots & \rho(k-2) \\ \rho(2) & \rho(1) & 1 & \cdots & \rho(k-3) \\ & & \cdots & & \\ \rho(k-1) & \rho(k-2) & & \cdots & 1 \end{bmatrix} \begin{bmatrix} \phi_{1k} \\ \phi_{2k} \\ \phi_{3k} \\ \cdots \\ \phi_{kk} \end{bmatrix} = \begin{bmatrix} \rho(1) \\ \rho(2) \\ \rho(3) \\ \cdots \\ \rho(k) \end{bmatrix}$$

we obtain the solution

$$\phi = R^{-1}\rho$$

How to compute PACF?

- Suppose $k = 1$, then
$$\phi_{11} = \rho(1)$$
- Suppose $k = 2$, then $R_{2\times 2}\phi_{2\times 1} = \rho_{2\times 1}$
$$\begin{bmatrix} 1 & \rho(1) \\ \rho(1) & 1 \end{bmatrix} \begin{bmatrix} \phi_{12} \\ \phi_{22} \end{bmatrix} = \begin{bmatrix} \rho(1) \\ \rho(2) \end{bmatrix}$$
- From $\phi = R^{-1}\rho$, we obtain
$$\begin{bmatrix} \phi_{12} \\ \phi_{22} \end{bmatrix} = \begin{bmatrix} \frac{\rho(1)[1-\rho(2)]}{1-\rho(1)^2} \\ \frac{\rho(2)[1-\rho(1)^2]}{1-\rho(1)^2} \end{bmatrix}$$

and thus
$$\phi_{22} = \frac{\rho(2) - \rho(1)^2}{1 - \rho(1)^2}$$

How to compute PACF?

- Suppose $k = 3$, then $R_{3\times 3}\phi_{3\times 1} = \rho_{3\times 1}$

$$\begin{bmatrix} 1 & \phi(1) & \phi(2) \\ \phi(1) & 1 & \phi(1) \\ \phi(2) & \phi(1) & 1 \end{bmatrix} \begin{bmatrix} \phi_{13} \\ \phi_{23} \\ \phi_{33} \end{bmatrix} = \begin{bmatrix} \rho(1) \\ \rho(2) \\ \rho(3) \end{bmatrix}$$

$$\Longrightarrow \quad \phi_{3\times 1} = R^{-1}_{3\times 3}\rho_{3\times 1}$$

- The last entry of ϕ gives ψ_{33}.

Special feature of PACF for AR process

- If $\{X_t\}$ is an AR(p) process, then PACF will be such that

$$\phi_{kk} = 0, \quad \text{for all } k > p.$$

- This feature characterizes the AR(p) process. We say PACF of AR(p) "cuts off" after lag p.

Special feature of PACF for AR process

- Reasoning: If $\{X_t\}$ is AR(p), then its ACF satistifies the YW equations of order p. If we solve the YW equations of order $k > p$, then the solution must be

$$\begin{aligned} \phi_{k\times 1} &= (\phi_{1k}, \phi_{2k}, \ldots, \phi_{(p+1)k}, \ldots, \phi_{kk})' \\ &= (\phi_{1k}, \phi_{2k}, \ldots, 0, \ldots, 0)' \end{aligned}$$

- Thus $\phi_{kk} = 0$ for $k > p$ and $\phi_{kk} = \phi_p$ for $k = p$.
- For example with AR(2)

$$\begin{aligned} \rho(1) &= \phi_1 + \phi_2\rho(1) + 0\rho(2) \\ \rho(2) &= \phi_1\rho(1) + \phi_2 + 0\rho(1) \\ \rho(3) &= \phi_1\rho(2) + \phi_2\rho(1) + 0 \end{aligned}$$

Interpretation of PACF

- Intuitively when solving the $k \times k$ system of YW equations, we are finding the theoretical coefficients in a vest fitting AR model of order k for the process.
- This is, consider

$$X_t = c_1 X_{t-1} + c_2 X_{t-2} + \cdots + c_k X_{t-k} + \epsilon_t$$

and find $c_1, c_2, \ldots, c_k$ to minimize the mean square error (MSE)

$$E\left[\left(X_t - (c_1 X_{t-1} + c_2 X_{t-2} + \cdots + c_k X_{t-k})\right)^2\right]$$

Interpretation of PACF

- The coefficients $c_1, c_2, \ldots, c_k$ that minimize the MSE are the same as solution to the $k \times k$ YW equations, because

$$c = \phi_{k\times 1} = R_{k\times k}^{-1} \rho_{k\times 1}$$

  where $c = (c_1, c_2, \ldots, c_k)'$.
- In fact, ϕ_{kk} is the "partial correlation" between the RVs X_t and X_{t-k} at lag k, after adjusting (or accounting for) the effects of the $(k-1)$ intermediate values $X_{t-1}, X_{t-2}, \cdots, X_{t-k+1}$.

Interpretation of PACF

- Consider least squares (LS) regression of X_t on $X_{t-1}, X_{t-2}, \cdots, X_{t-k+1}$. The fitted values are

$$\begin{aligned}\hat{X}_t &= c_1 X_{t-1} + c_2 X_{t-2} + \cdots + c_{k-1} X_{t-k+1} \\ &= \phi_{1,k-1} X_{t-1} + \phi_{2,k-1} X_{t-2} + \cdots + \phi_{k-1,k-1} X_{t-k+1}\end{aligned}$$

  with "residuals" (or adjusted values)

$$X_t^* = X_t - \hat{X}_t$$

- Then

$$\phi_{kk} = \mathrm{Cor}(Y_t^*, Y_{t-k}^*)$$

  is the ordinary correlation between X_t^* and X_{t-k}^*.

Interpretation of PACF

For example, with $k = 2$, $\phi_{11} = \rho(1)$ and

$$X_t^* = X_t - \phi_{11} X_{t-1}, \qquad X_{t-2}^* = X_{t-2} - \phi_{11} X_{t-1}$$

we have

$$\begin{aligned}
\phi_{22} &= \mathrm{Cor}(X_t^*, X_{t-2}^*) \\
&= \mathrm{Cor}(X_t - \phi_{11} X_{t-1}, X_{t-2} - \phi_{11} X_{t-1}) \\
&= \frac{\mathrm{Cor}(X_t - \phi_{11} X_{t-1}, X_{t-2} - \phi_{11} X_{t-1})}{\sqrt{\mathrm{Var}(X_t - \phi_{11} X_{t-1})}\sqrt{\mathrm{Var}(X_{t-2} - \phi_{11} X_{t-1})}} \\
&= \frac{\rho(2) - \rho(1)^2}{1 - \rho(1)^2}
\end{aligned}$$

Sample PACF

For a sample of T observations $X_1, X_2, \ldots, X_t$,

1. Obtain sample ACF values

$$\hat{\rho}(1), \hat{\rho}(2), \ldots, \hat{\rho}(k)$$

2. Obtain solution to the $k \times k$ system of sample YW equations

$$\hat{\phi}_{k\times 1} = \hat{R}^{-1}_{k\times k}\hat{\rho}_{k\times 1}$$

3. The last entry of $\hat{\phi}_{k\times 1}$ is $\hat{\phi}_{kk}$, which is the sample PACF at lag k.
   - If $k = 1$,

$$\hat{\phi}_{11} = \hat{\rho}(1)$$

   - If $k = 2$,

$$\hat{\phi}_{22} = \frac{\hat{\rho}(2) - \hat{\rho}(1)^2}{1 - \hat{\rho}(1)^2}$$

Invertibility of MA(q)

- Consider an MA(q) model

$$X_t = \mu + \epsilon_t - \sum_{i=1}^{q} \theta_i \epsilon_{t-i} = \mu + \theta(B)\epsilon_t$$

- For certain purposes such as forecasting, it is useful to represent the model in infinite AR form as

$$X_t = \sum_{j=1}^{\infty} \pi_j X_{t-j} + \delta + \epsilon_t$$

- An MA(q) model is **invertible** if it can be expressed in infinite AR form as above with $\sum_{j=1}^{\infty} |\pi_j| < \infty$.

Invertibility of MA(q)

- To obtain the infinite AR form, we apply $\theta(B)^{-1}$ to both sides of the MA(q) equation

$$\theta(B)^{-1}X_t = \theta(B)^{-1}\mu + \epsilon_t$$

- That is,

$$\pi(B)X_t = \delta + \epsilon_t$$

where

$$\pi(B) = \theta(B)^{-1} = 1 - \sum_{j=1}^{\infty} \pi_j \mathbf{B}^j$$

is the infinite AR operator.

Invertibility of MA(q)

- The coefficients $\{\pi_j\}$ are determined by the relation

$$\theta(B)\pi(B) = 1$$

- That is

$$\begin{aligned} 1 &= (1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q)(1 - \pi_1 B - \pi_2 B^2 - \cdots) \\ &= 1 - (\pi_1 + \theta_1)B - (\pi_2 + \theta_1\pi_1 + \theta_2)B^2 - \cdots \\ &\quad - (\pi_j + \theta_1\pi_{j-1} - \cdots - \theta_q\pi_{j-q}B^j) - \cdots \end{aligned}$$

- It follows that, for $j = 1, 2, \ldots,$

$$\pi_j - \theta_1\pi_{j-1} - \theta_2\pi_{j-2} - \cdots - \theta_q\pi_{j-q} = 0$$

- Also with conventions

$$\pi_0 = -1, \quad \pi_j = 0, \quad j < 0$$

Invertibility of MA(q)

- Valid infinite representation exists only if $\sum_{j=1}^{\infty} |\pi_j| < \infty$.
- By theory of solutions to difference equations, explicit form of π_j is determined by q roots $m_1, m_2, \ldots, m_q$ of

$$m^q - \theta_1 m^{q-1} \theta_2 m^{q-2} - \cdots - \theta_q = 0$$

- It can be shown that $\sum_{j=1}^{\infty} |\pi_j| < \infty$ if and only if $|m_j| < 1$ for all $j = 1, 2, \ldots, q$.
- Condition for invertibility of MA(q): All roots $m_1, m_2, \ldots, m_q$ of $m^q - \theta_1 m^{q-1} - \theta_2 m^{q-2} - \cdots - \theta_q = 0$ are less than 1 in absolute value.

Invertibility of MA(q)

- Another reason for considering invertibility for MA(q) is to obtain a unique MA(q) model.
- That is, for any given MA(q) model with $\gamma(0), \rho(1), \ldots, \rho(q)$, there are several MA($q$) model equations (up to 2^q) which give rise to the same values, so that the set of conefficients $\theta_1, \theta_2, \ldots, \theta_q$ is not unique.
- For example, MA(2), there are up to $2^2 = 4$ equivalent versions of the model.
- Among all the different equivalent models, only one version satisfies the invertibility condition.
- Thus by convention, we always "choose" the inverible MA(q) model for uniqueness.

Least-squares estimator (LSE) of AR(1)

- We will find the OLS estimator of the coefficient ϕ of AR(1) with mean zero (If $\mu \neq 0$, take $X_t - \mu$):

$$X_t = \phi X_{t-1} + \epsilon_t \quad \{\epsilon_t\} \sim \text{iid } (0, \sigma_\epsilon^2)$$

- Sum of squared errors from data $X_1, X_2, \cdots, X_n$:

$$S(\phi) = \sum_{t=2}^{n} (X_t - \phi X_{t-1})^2.$$

- We will find ϕ so that $S(\phi)$ is minimized and the OLS estimator of AR(1) is given by

$$\hat{\phi} = \frac{\sum_{t=2}^{n} X_t X_{t-1}}{\sum_{t=2}^{n} X_{t-1}^2}$$

Central Limit Theorem (CLT) of AR(1):
In AR(1), $X_t = \phi X_{t-1} + \epsilon_t$, with $\{\epsilon_t\} \sim$ iid $(0, \sigma_\epsilon^2)$, the OLS estimator $\hat{\phi}$ follows the asymptotic normality as $n \to \infty$:

$$\sqrt{n}\,(\hat{\phi} - \phi) \xrightarrow{d} N\left(0, \frac{\sigma_\epsilon^2}{EX_1^2}\right)$$

Forecasting in AR(1)

- Conditional expectation Y given X: $E[Y|X]$ is a function of X, because
$$E[Y|X = x] = \int yf(y|x)dy$$
which is a function of x.
- Conditional expectation Y given $X_1, X_2 \cdots, X_t$: $E[Y|X_1, \cdots, X_t]$ is a function of $(X_1, X_2, \cdots, X_t)$.
- $E[g(X_1, X_2 \cdots, X_n)|X_1, \cdots, X_n] = g(X_1, X_2 \cdots, X_n)$ where g is a function from $\mathbb{R}^n$ to $\mathbb{R}$.

Forecasting in AR(1)

Now we will forecast future data. Let t be the present time. Suppose that we have information $X_1, \cdots, X_t$ and mean μ, coefficient ϕ.

One-step ahead forecast:

- Let $\hat{X}_{t+1} \equiv \hat{X}_t(1)$ be one-step ahead forecast of AR(1) model.

$$\begin{aligned}
\hat{X}_{t+1} = \hat{X}_t(1) &= E[X_{t+1}|X_1, \cdots, X_t] \\
&= E[\mu + \phi(X_t - \mu) + \epsilon_{t+1}|X_1, \cdots, X_t] \\
&= \mu + \phi(X_t - \mu) + E[\epsilon_{t+1}|X_1, \cdots, X_t] \\
&= \mu + \phi(X_t - \mu).
\end{aligned}$$

Two-step ahead forecast:

- Let $\hat{X}_{t+2} \equiv \hat{X}_t(2)$ be two-step ahead forecast of AR(1) model.

$$\begin{aligned}
\hat{X}_{t+2} = \hat{X}_t(2) &= E[X_{t+2}|X_1, \cdots, X_t] \\
&= E[\mu + \phi(X_{t+1} - \mu) + \epsilon_{t+2}|X_1, \cdots, X_t] \\
&= \mu + \phi(E[X_{t+1}|X_1, \cdots, X_t] - \mu) + E[\epsilon_{t+2}|X_1, \cdots, X_t] \\
&= \mu + \phi(\hat{X}_t(1) - \mu) + 0 \\
&= \mu + \phi(\mu + \phi(X_t - \mu) - \mu) \\
&= \mu + \phi^2(X_t - \mu)
\end{aligned}$$

Forecasting in AR(1)

- By the mathematical induction, we may assume

$$\hat{X}_{t+\ell-1} = \hat{X}_t(\ell-1) = \mu + \phi(\hat{X}_t(\ell-2) - \mu)$$

- Note that

$$\hat{X}_t(\ell-1) - \mu = \phi(\hat{X}_t(\ell-2) - \mu)$$

and for each $k = 2, 3, \cdots, \ell - 1$,

$$\hat{X}_t(k) - \mu = \phi(\hat{X}_t(k-\ell) - \mu)$$

Forecasting in AR(1) model

ℓ-step ahead forecast:

- Let $\hat{X}_{t+\ell} \equiv \hat{X}_t(\ell)$ be ℓ-step ahead forecast of AR(1) model.

$$\begin{aligned}
\hat{X}_{t+\ell} = \hat{X}_t(\ell) &= E[X_{t+\ell}|X_1, \cdots, X_t] \\
&= E[\mu + \phi(X_{t+\ell-1} - \mu) + \epsilon_{t+\ell}|X_1, \cdots, X_t] \\
&= \mu + \phi(\hat{X}_t(\ell-1) - \mu) \\
&= \mu + \phi^2(\hat{X}_t(\ell-2) - \mu) \\
&\cdots \\
&= \mu + \phi^{\ell-1}(\hat{X}_t(1) - \mu) \\
&= \mu + \phi^t(X_t - \mu)
\end{aligned}$$

- Since $|\phi| < 1$, $\phi^\ell \to 0$ as $\ell \to \infty$. Thus, $\hat{X}_t(\ell) \to \mu$ as $\ell \to \infty$.
- It means that future values of stationary AR(1) approaches to the mean as time goes to infinity.

Outline

Autoregressive Moving Average (ARMA) model

- A process $\{X_t\}$ is autoregressive moving average of order (p, q), denoted as ARMA(p, q), if it is stationary and

$$X_t - \sum_{i=1}^{p} \phi_i X_{t-i} = \delta + \epsilon_t - \sum_{i=1}^{q} \theta_i \epsilon_{t-i}$$

  where $\{\epsilon_t\}$ is a white noise process (i.e. ϵ_t are independent RVs with mean 0 and variance σ^2 for all t).
- The parameters p and q are called the autoregressive and the moving average orders, respectively.

Autoregressive Moving Average (ARMA) model

- If X_t has a nonzero mean δ, we set $\delta = \mu(1 - \phi_1 - ... - \phi_p)$ and write the model as

$$(X_t - \mu) - \sum_{i=1}^{p} \phi_i (X_t - \mu) = \epsilon_t - \sum_{i=1}^{q} \theta_i \epsilon_{t-i}$$

- In an operator form, ARMA(p, q) can be written as,

$$\phi(B) X_t = \delta + \theta(B) \epsilon_t$$

where

$$\phi(B) = 1 - \sum_{i=1}^{p} \phi_i B^i, \quad \theta(B) = 1 - \sum_{i=1}^{q} \theta_i B^i$$

Stationarity of ARMA process

- **Condition for stationarity of ARMA(p, q):** All roots $m_1, m_2, \ldots, m_p$ of

$$m^p - \phi_1 m^{p-1} - \phi_2 m^{p-2} - \cdots - \phi_p = 0$$

are all less than 1 in absolute value.

- A stationary ARMA(p, q) process can be expressed in MA(∞):

$$X_t = \phi(B)^{-1}\delta + \phi(B)^{-1}\theta(B)\epsilon_t = \mu + \psi(B)\epsilon_t$$

where

$$\psi(B) = \phi(B)^{-1}\theta(B) = \sum_{j=1}^{\infty} \psi_j R^j$$

with $\sum_{j=0}^{\infty} |\psi_j| < \infty$.

- It is also called **causality** of ARMA processes.

Invertibility of ARMA process

- **Condition for invertibility of ARMA(p, q):** All roots $m_1, m_2, \ldots, m_q$ of

$$m_q - \theta_1 m^{q-1} - \theta_2 m^{q-2} - \cdots - \theta_q = 0$$

are less than 1 in absolute value.

Invertibility of ARMA process

- An invertible ARMA(p, q) process can be expressed in AR(∞):

$$\theta(B)^{-1}\phi(B)X_t = \theta(B)^{-1}\delta + \epsilon_t$$

- Equivalently, it can be rewritten as

$$\pi(B)X_t = \delta^* + \epsilon_t$$

where

$$\pi(B) = \theta(B)^{-1}\phi(B) = 1 - \sum_{j=1}^{\infty} \pi_j B^j$$

with $\sum_{j=1}^{\infty} |\pi_j| < \infty$.

Mean of ARMA process

- Let $\{X_t\}$ denote a stationary ARMA(p, q) process with the equation

$$X_t - \sum_{i=1}^{p} \phi_i X_{t-i} = \delta + \epsilon_t - \sum_{i=1}^{p} \theta_i \epsilon_{t-i}$$

- Thus

$$\mu = E(X_t) = \frac{\delta}{1 - \sum_{i=1}^{p} \phi_i}$$

Autocovariance function

- For $0 \leq k \leq q$,

$$\gamma(k) = \sum_{i=1}^{p} \phi_i \gamma(k-i) - \sigma^2(\theta_j \psi_0 + \theta_{k+1}\psi_1 + \cdots + \theta_q \psi_{q-k})$$

  with the convention that $\theta_0 = -1$.
- Thus for $k = 0$,

$$\gamma(0) = \mathrm{Var}(X_t) = \sum_{i=1}^{p} \phi_i \gamma(i) + \sigma^2(1 - \sum_{i=1}^{q} \theta_i \psi_i)$$

- For $k > q$,

$$\gamma(k) = \sum_{i=1}^{p} \phi_i \gamma(k-i)$$

Autocorrelation function

- Note the infinite MA form gives

$$X_{t-k} = \mu + \sum_{j=0}^{\infty} \psi_j \epsilon_{t-k-j}$$

and thus

$$\mathrm{Cov}(X_{t-k}, \epsilon_{t-i}) = \begin{cases} \psi_{j-k}\sigma^2 & \text{if } i-k \geq 0 \\ 0 & \text{if } i-k < 0 \end{cases}$$

- The generalized Yule-Walker equations are

$$\rho(k) = \sum_{j=1}^{p} \phi_i \rho(k-i), \quad \text{for } k > q$$

- These equations can be used to determine the autocovariance function any time lag for any ARMA(p, q).

ARMA(1,1) process

- With $p = q = 1$, we have the ARMA(1,1) equation

$$(1 - \phi B)X_t = (1 - \theta B)\epsilon_t$$

- Stationarity and infinite MA form: ARMA(1,1) is stationary if $|\phi| < 1$. Then

$$\psi_j - \phi\psi_{j-1} = \begin{cases} -\theta & j = 1 \\ 0 & j > 1 \end{cases}$$

and $\psi_0 = 1$.
- In fact, $\{\psi_j\}$ exponentially decline with rate ϕ,

$$\psi_j = \phi^j\Big(1 - \frac{\theta}{\phi}\Big), \quad \text{for } j > 0$$

ARMA(1,1) process

- Invertibility and infinite AR form: ARMA(1,1) is invertible if $|\theta| < 1$.
- Then

$$\pi_j - \theta\pi_{j-1} = \begin{cases} \phi & j = 1 \\ 0 & j > 1 \end{cases}$$

and $\pi_0 = -1$.
- In fact, $\{\pi_j\}$ exponentially declines with rate θ,

$$\pi_j = \theta^j\Big(\frac{\phi}{\theta} - 1\Big), \quad \text{for } j > 0$$

ARMA(1,1) process

- Autocovariances of ARMA(1,1) are determined by

$$\begin{aligned}
\gamma(0) &= \phi\gamma(1) + \sigma^2(1 - \theta\psi_1) \\
\gamma(1) &= \phi\gamma(0) + \sigma^2(\theta\psi_0) \\
\gamma(k) &= \phi\gamma(k-1), \quad k > 1
\end{aligned}$$

  where $\psi_1 = \phi - \theta, \psi_0 = 1$.
- The first two equations are used to solve for $\gamma(0)$ and $\gamma(1)$

$$\gamma(0) = \sigma^2 \left(\frac{1 - 2\phi\theta + \theta^2}{1 - \phi^2} \right), \quad \rho(1) = \frac{(1 - \phi\theta)(\phi - \theta)}{1 - 2\phi\theta + \theta^2}$$

- The rest of ACF values of ARMA(1,1) are according to the generalized Yule-Walker equations

$$\rho(k) = \phi\rho(k-1), \quad k = 2, 3, \ldots$$

ARMA(1,1) example

- Consider ARMA(1,1) with $\phi = 0.8, \theta = -0.6$.
- Stationarity and infinite MA form.
- Invertibility and infinite AR form.
- Mean, autocovariances, ACF, and PACF.
- The ACF values are

$$\rho(1) = 0.893, \rho(2) = 0.714, \rho(3) = 0.572,$$
$$\rho(4) = 0.457, \rho(5) = 0.366$$

- The PACF values are

$$\phi_{11} = 0.893, \phi_{22} = -0.411, \phi_{33} = 0.227,$$
$$\phi_{44} = -0.133, \phi_{55} = 0.079.$$

ARMA(2,1) example

- For an example, consider an ARMA(2,1) model

$$(1 - 0.5B^2)X_t = (1 + 0.25B)\epsilon_t$$

  - The equation $\phi(x) = 1 - 0.5x^2 = 0$ has roots $x = \pm\sqrt{2}$. Since $|x| > 1$, the model is stationary.
  - The equation $\theta(x) = 1 + 0.25x = 0$ has roots $x = -4$. Since $|x| > 1$, the model is invertible.

- Therefore, it can be written as AR(∞), $X_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$, by solving

$$X_t = \frac{1 + 0.25B}{1 - 0.5B^2}\epsilon_t = \sum_{j=0}^{\infty} \psi_j B^j \epsilon_t$$

ARMA(2,2) example

- Consider the following ARMA(2,2) model:

$$X_t = 0.4X_{t-1} + 0.45X_{t-2} + \epsilon_t + \epsilon_{t-1} + 0.25\epsilon_{t-2}.$$

- Using the backshift operator B,

$$(1 - 0.4B + 0.45B^2)X_t = (1 + B + 0.25B^2)\epsilon_t$$

- Note that
  - $\phi(x) = 1 - 0.4x - 0.45x^2 = (1 + 0.5x)(1 - 0.9x)$
  - $\theta(x) = 1 + x + 0.25x^2 = (1 + 0.5x)^2$

ARMA(2,2) example

- All roots are outside of the unit circle on the complex plain, and thus the model is stationary and invertible.
- Hence,

$$X_t = \frac{\theta(B)}{\phi(B)}\epsilon_t = \frac{(1+0.5B)^2}{(1+0.5B)(1-0.9B)}\epsilon_t = \frac{1+0.5B}{1-0.9B}\epsilon_t.$$

- Therefore, the model is in fact an ARMA(1,1) model:

$$(1-0.9B)X_t = (1+0.5B)\epsilon_t, \quad X_t = 0.9X_{t-1} + \epsilon_t + 0.5\epsilon_{t-1}.$$

- This is a parameter redundancy problem.

Forecasting in ARMA(1,1)

- We study forecasting future values of a stationary ARMA(1,1) model with mean μ.

$$X_t = \mu + \phi(X_{t-1} - \mu) + \epsilon_t + \theta\epsilon_{t-1}, \quad |\phi| < 1, \quad |\theta| < 1.$$

- Suppose that we have information $X_1, X_2, \cdots, X_t$ and mean μ, coefficients ϕ and θ.
- Since $|\theta| < 1$, the model is invertible, and thus ϵ_t can be expressed as a linear combination of $X_1, X_2, \cdots, X_t$ and also X_t can be expressed as a linear combination of $\epsilon_1, \cdots, \epsilon_t$.
- Therefore, we regard the condition given $\{X_1, X_2, \cdots, X_t\}$ as the condition given $\{\epsilon_1, \cdots, \epsilon_t\}$:

$$E[Y|X_1, X_2, \cdots, X_t] = E[Y|\epsilon_1, \cdots, \epsilon_t].$$

Forecasting in ARMA(1,1)

One-step ahead forecast:

- Let $\hat{X}_{t+1} \equiv \hat{X}_t(1)$ be the one-step ahead forecast of ARMA(1,1).

$$\begin{aligned}
\hat{X}_{t+1} &= \hat{X}_t(1) \\
&= E[X_{t+1}|X_1, \cdots, X_t] \\
&= E[\mu + \phi(X_t - \mu) + \epsilon_{t+1} + \theta\epsilon_t|X_1, \cdots, X_t] \\
&= \mu + \phi(X_t - \mu) + E[\epsilon_{t+1}|X_1, \cdots, X_t] + \theta E[\epsilon_t|X_1, \cdots, X_t] \\
&= \mu + \phi(X_t - \mu) + \theta\epsilon_t.
\end{aligned}$$

Forecasting in ARMA(1,1)

Two-step ahead forecast:

- Let $\hat{X}_{t+2} \equiv \hat{X}_t(2)$ be the two-step ahead forecast of ARMA(1,1).

$$\begin{aligned}
\hat{X}_{t+2} = \hat{X}_t(2) &= E[X_{t+2}|X_1, \cdots, X_t] \\
&= E[\mu + \phi(X_{t+1} - \mu) + \epsilon_{t+2} + \theta\epsilon_{t+1}|X_1, \cdots, X_t] \\
&= \mu + \phi(E[X_{t+1}|X_1, \cdots, X_t] - \mu) \\
&= \mu + \phi(\hat{X}_t(1) - \mu) \\
&= \mu + \phi[(\mu + \phi(X_t - \mu)\theta\epsilon_t) - \mu] \\
&= \mu + \phi^2(X_t - \mu) + \phi\theta\epsilon_t
\end{aligned}$$

- We may assume that, for some $\ell \geq 2$,

$$\hat{X}_{t+\ell-1} = \hat{X}_t(\ell - 1) = \mu + \phi^{\ell-1}(X_t - \mu) + \theta\phi^{\ell-2}\epsilon_t.$$

Forecasting in ARMA(1,1)

ℓ-step ahead forecast:

- Let $\hat{X}_{t+\ell} \equiv \hat{X}_t(\ell)$ be the ℓ-step ahead forecast of ARMA(1,1).

$$\begin{aligned}
\hat{X}_{t+\ell} = \hat{X}_t(\ell) &= E[X_{t+\ell}|X_1, \cdots, X_t] \\
&= E[\mu + \phi(X_{t+\ell-1} - \mu) + \epsilon_{t+\ell} + \theta\epsilon_{t+\ell-1}|X_1, \cdots, X_t] \\
&= \mu + \phi(\hat{X}_t(\ell-1) - \mu) + 0 + 0 \\
&= \mu + \phi[\phi^{\ell-1}(X_t - \mu) + \theta\phi^{\ell-2}\epsilon_t] \\
&= \mu + \phi^{\ell}(X_t - \mu) + \theta\phi^{\ell-1}\epsilon_t
\end{aligned}$$

- Since $|\phi| < 1$, $\phi^{\ell} \to 0$ as $\ell \to \infty$. Thus, $\hat{X}_t(\ell) \to \mu$ as $\ell \to \infty$.
- It means that future values of stationary ARMA(1,1) approaches to the mean as time goes to infinity.

Outline

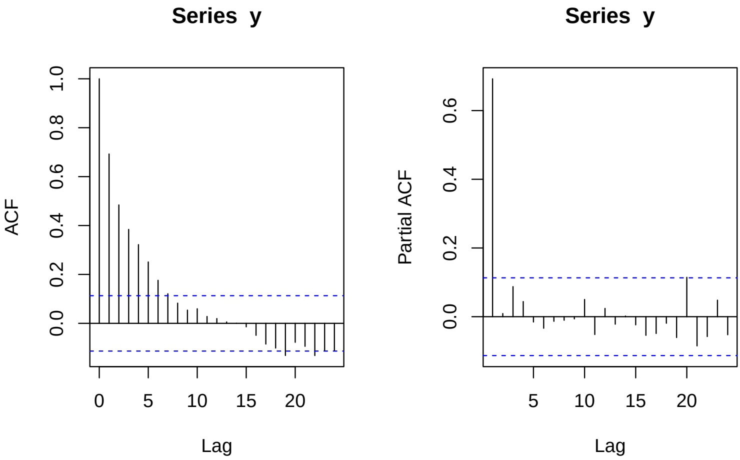
Behavior of ACF and PACF for ARMA

Table 3.1. Behavior of the ACF and PACF for ARMA models

	AR(p)	MA(q)	ARMA(p, q)
ACF	Tails off	Cuts off after lag q	Tails off
PACF	Cuts off after lag p	Tails off	Tails off

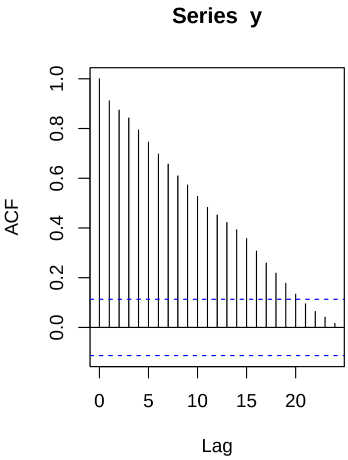
AR(1) example

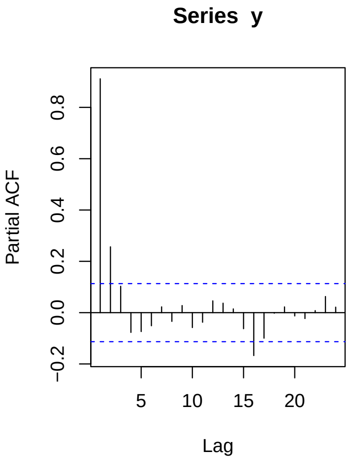
```
y = arima.sim(n=300, model=list(ar=0.7))
par(mfrow=c(1,2))
acf(y); pacf(y)
```



AR(2) example

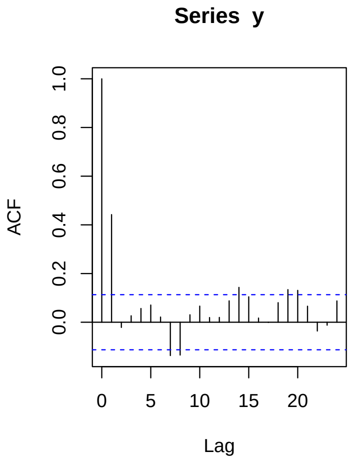
```
y = arima.sim(n=300, model=list(ar=c(0.7,0.2)))
par(mfrow=c(1,2))
acf(y); pacf(y)
```

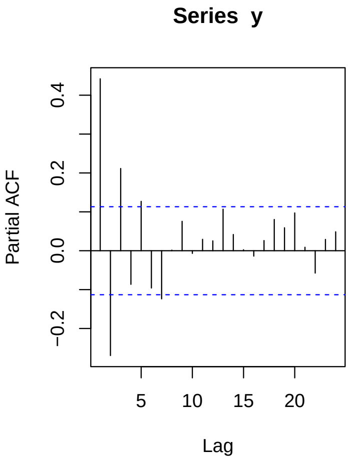




MA(1) example

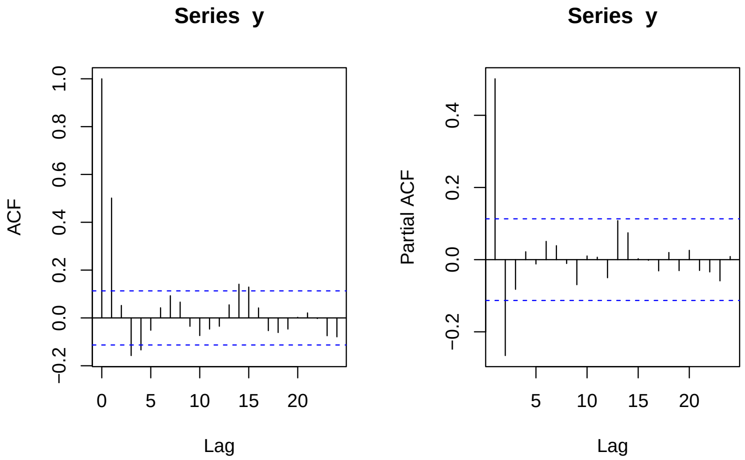
```
y = arima.sim(n=300, model=list(ma=0.7))
par(mfrow=c(1,2))
acf(y); pacf(y)
```





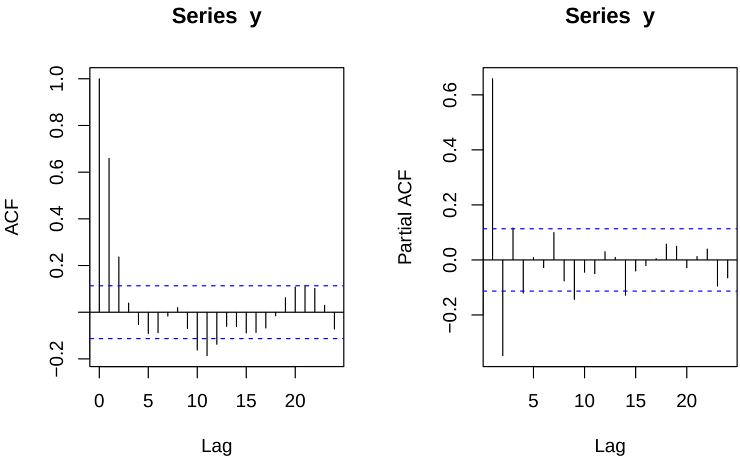
MA(2) example

```r
y = arima.sim(n=300, model=list(ma=c(0.7,0.2)))
par(mfrow=c(1,2))
acf(y); pacf(y)
```



ARMA(1,1) example

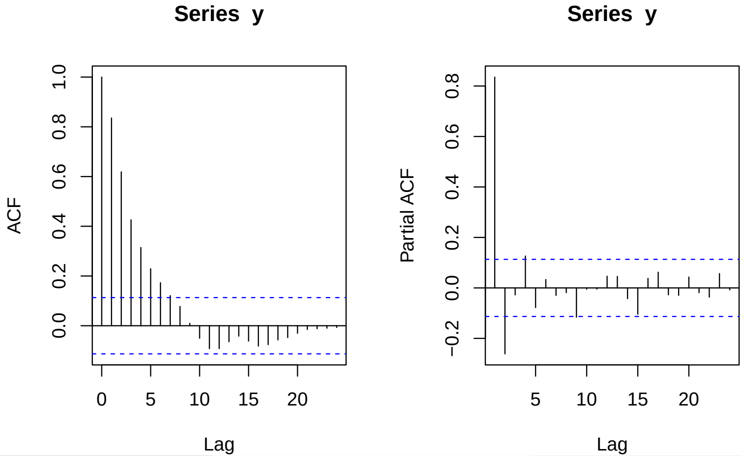
```
y = arima.sim(n=300, model=list(ar=0.5,ma=0.5))
par(mfrow=c(1,2))
acf(y); pacf(y)
```



ARMA(2,2) example

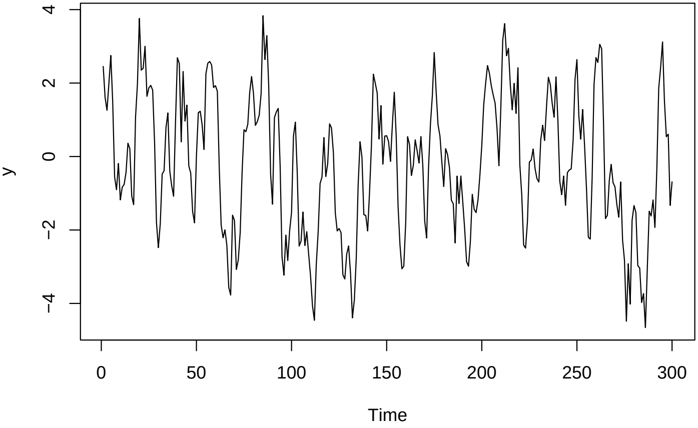
Note that ARMA(2,2) is a stationary process.

```
y = arima.sim(n=300, model=list(ar=c(0.5,0.2),ma=c(0.5,0.2)))
par(mfrow=c(1,2))
acf(y); pacf(y)
```



Fitting an ARMA model

```r
ts.plot(y)
```



ARMA fitting and model selection

```
Arima(y, order=c(1,0,0))$aic #AR(1)
## [1] 865.5338
Arima(y, order=c(2,0,0))$aic #AR(2)
## [1] 845.1778
Arima(y, order=c(0,0,1))$aic #MA(1)
## [1] 1009.244
Arima(y, order=c(0,0,2))$aic #MA(2)
## [1] 888.5498
Arima(y, order=c(1,0,1))$aic #ARMA(1,1)
## [1] 849.4444
Arima(y, order=c(2,0,1))$aic #ARMA(2,1)
## [1] 847.0813
Arima(y, order=c(2,0,2))$aic #ARMA(2,2)
## [1] 841.7542
Arima(y, order=c(3,0,3))$aic #ARMA(3,3)
## [1] 838.7369
```

ARMA fitting and model selection

```r
library(forecast)
auto.arima(y)
## Series: y
## ARIMA(1,0,2) with zero mean
##
## Coefficients:
##          ar1     ma1     ma2
##       0.6457  0.4392  0.2490
## s.e.  0.0684  0.0850  0.0759
##
## sigma^2 = 0.9421:  log likelihood = -415.95
## AIC=839.9   AICc=840.04   BIC=854.72
```