Chapter 2: Stationary ARMA processes

Jesús Bueren

EUI

Introduction

- This chapter follows chapter 3 in Hamilton.
- It provides a class of models for describing the dynamics of an individual time series.
- We first go through a set of basic time series concepts and the properties of various ARMA processes.

Definitions

Ensemble mean

- Imagine a sequence of I independent computers generating sequences of random numbers from a distribution with finite first and second moments:

$$\{y_t^{(1)}\}_{t=-\infty}^{\infty}; \{y_t^{(2)}\}_{t=-\infty}^{\infty}; \cdots ; \{y_t^{(I)}\}_{t=-\infty}^{\infty}$$

$y_t^{(i)}$ is a draw from the random variable Y_t

- The ensemble mean is defined as:

$$E[Y_t] = \int_{-\infty}^{\infty} y_t f_{Y_t}(y_t) dy_t = \underset{I \to \infty}{\text{plim}} (1/I) \sum_{i=1}^{I} y_t^{(i)} = \mu_t$$

Definitions

Autocovariance

- The autocovariance is defined as:

$$
\begin{aligned}
&E[(Y_t - \mu_t)(Y_{t-j} - \mu_{t-j})] \\
&\qquad = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y_t - \mu_t)(y_{t-j} - \mu_{t-j}) f_{Y_t, Y_{t-j}}(y_t, y_{t-j}) dy_t dy_{t-j} \\
&\qquad = \operatorname*{plim}_{I \to \infty} (1/I) \sum_{i=1}^{I} (y_t^{(i)} - \mu_t)(y_{t-j}^{(i)} - \mu_{t-j}) = \gamma_{jt}
\end{aligned}
$$

Definitions

Stationarity

- If neither the mean, nor the autocovariances depend on date t, then the process Y_t is said to be covariance-stationary or weakly stationary.
    - $E[Y_t] = \mu \ \forall \ t$
    - $E[(Y_t - \mu)(Y_{t-j} - \mu)] = \gamma_j \ \forall \ t$

Definitions

Ergodicity

- A stationary process is said to be ergodic if:

$$\underset{T\to\infty}{\text{plim}}\, 1/T \sum_{t=1}^{T} y_t^{(i)} = \underset{I\to\infty}{\text{plim}}\, 1/I \sum_{i=1}^{I} y_t^{(i)} = \mu$$

- Example of a non-ergodic stationary process:

$$y_t^{(i)} = \mu^{(i)} + \epsilon_t; \mu^{(i)} \sim N(0, \lambda); \epsilon_t \sim N(0, \sigma)$$

- Sufficient conditions for ergodicity of a stationary process: $\sum_{j=0}^{\infty} |\gamma_j| < \infty$

Moving-Average Processes
MA(1)

- Let $\{\epsilon_t\}$, $\epsilon_t \sim N(0, \sigma^2)$, i.i.d: Gaussian white noise
- Consider the process:

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

this time series is called a *first-order moving average process*,denoted MA(1).

Moving Average Processes

MA(1)

- Expectation: $E[Y_t] = \mu$
- Autocovariance:

$$E[(Y_t - \mu)(Y_{t-j} - \mu)] = \begin{cases} \sigma^2(1+\theta^2), & \text{if } j = 0 \\ \theta\sigma^2, & \text{if } j = 1 \\ 0, & \text{otherwise} \end{cases}$$

$\Rightarrow$ Stationary

- $\sum_{j=0}^{\infty} |\gamma_j| = \sigma^2(1+\theta^2) + |\theta|\sigma^2$
  $\Rightarrow$ Ergodic

Moving Average Processes
MA(q)

- Expectation: $E[Y_t] = \mu$
- Autocovariance:

$$E[(Y_t - \mu)(Y_{t-j} - \mu)] = \begin{cases} \sigma^2(1 + \sum_{i=1}^{q} \theta_i^2), & \text{if } j = 0 \\ \sigma^2(\theta_j + \sum_{i=1}^{q-j} \theta_i \theta_{i+j}), & \text{if } 0 < j <= q \\ 0, & \text{otherwise} \end{cases}$$

$\Rightarrow$ Stationary

- $\sum_{j=0}^{\infty} |\gamma_j| < \infty$
  $\Rightarrow$ Ergodic

Moving Average Processes
MA(∞)

- Expectation: $E[Y_t] = \mu$
- Autocovariance:

$$E[(Y_t - \mu)(Y_{t-j} - \mu)] = \begin{cases} \sigma^2(1 + \sum_{i=1}^{\infty} \theta_i^2), & \text{if } j = 0 \\ \sigma^2(\theta_j + \sum_{i=1}^{\infty} \theta_i \theta_{i+j}), & \text{if } j > 0 \end{cases}$$

$\Rightarrow$ Stationary

- $\sum_{j=0}^{\infty} |\gamma_j| < \infty$ if $\sum_{i=1}^{\infty} |\theta_i| < \infty$
$\Rightarrow$ Ergodic

Autoregressive Processes

AR(1)

- Let $\{\epsilon_t\}$, $\epsilon_t \sim N(0, \sigma^2)$, i.i.d: Gaussian white noise
- Consider the process:

$$Y_t = c + \phi Y_{t-1} + \epsilon_t,$$

this time series is called a *first-order autoregressive process*,denoted AR(1).

- Notice that this process takes the form of a first-order difference equation.
- We know from our analysis of first-order difference equations that if $|\phi| > 1$, the consequences of ϵ's for Y accumulate $\Rightarrow$ not covariance stationary

Autoregressive Processes

AR(1)

- The solution is given by:

$$\begin{aligned} Y_t =& (c + \epsilon_t) + \phi(c + \epsilon_{t-1}) + \phi^2(c + \epsilon_{t-1}) + \cdots \\ =& c/(1-\phi) + \epsilon_t + \phi\epsilon_{t-1} + \phi^2\epsilon_{t-2} + \cdots \end{aligned}$$

- This can be viewed as an MA(∞) process.
- With $|\phi| < 1$, $\sum_{i=1}^{\infty} |\phi^i| = 1/(1-|\phi|) < \infty \Rightarrow$ Ergodic.
- Autocovariance:

$$E[(Y_t - \mu)(Y_{t-j} - \mu)] = \sigma^2\phi^j/(1-\phi^2)$$

Autoregressive Processes

AR(2)

- A second-order autoregression AR(2) satisfies,

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \epsilon_t \tag{1}$$

  or in lag operation notation,

$$(1 - \phi_1 L - \phi_2 L^2) Y_t = \epsilon_t$$

- The process is stationary provided that the roots z_1 and z_2 of

$$1 - \phi_1 z - \phi_2 z^2 = 0$$

  lie outside the unit circle (or λ_1 and λ_2 smaller than one in modulus).

- We obtain:

$$(1 - \phi_1 L - \phi_2 L^2) = (1 - \lambda_1 L)(1 - \lambda_2 L),$$

  where $\lambda_1 = 1/z_1$ and $\lambda_2 = 1/z_2$

Autoregressive Processes

AR(2)

- To find autocovariances subtract the unconditional mean ($\mu = c/(1 - \phi_1 - \phi_2)$) on both sides of equation (1) multiply by $Y_{t-j} - \mu$ and take expectations:

$$\gamma_j = \phi_1 \gamma_{j-1} + \phi_2 \gamma_{j-2} \text{ for } j > 0 \tag{2}$$

- For the first 3 autocovariances we have:

$$\begin{aligned}
\gamma_0 =& \phi_1 \gamma_1 + \phi_2 \gamma_2 + \sigma^2 \\
\gamma_1 =& \phi_1 \gamma_0 + \phi_2 \gamma_1 \\
\gamma_2 =& \phi_1 \gamma_1 + \phi_2 \gamma_0
\end{aligned}$$

  which is a system of equations with 3 equations and 3 unknowns.

- For further autocovariances, iterate on equation (2).

Autoregressive Processes

AR(p)

- These techniques generalize in a straightforward way to pth-order difference equation of the form

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \epsilon_t \tag{3}$$

written in terms of the lag operator as:

$$(1 - \phi_1 L - \cdots - \phi_p L^p) y_t = \epsilon_t$$

- The process is stationary as long as the roots of:

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

lie outside the unit circle.

- Then,

$$(1 - \phi_1 L - \cdots - \phi_p L^p) = (1 - \lambda_1 L)(1 - \lambda_2 L)...(1 - \lambda_p L)$$

Autoregressive Processes

AR(p)

- To find autocovariances subtract the unconditional mean ($\mu = 1/(1 - \phi_1 - \dots \phi_p)$) on both sides of equation (3) multiply by $Y_{t-j} - \mu$ and take expectations:

$$\gamma_j = \phi_1 \gamma_{j-1} + \dots + \phi_p \gamma_{j-p} \tag{4}$$

- For the first p autocovariances we have:

$$\begin{aligned} \gamma_0 =& \phi_1 \gamma_1 + \dots + \phi_p \gamma_p + \sigma^2 \\ \gamma_1 =& \phi_1 \gamma_0 + \dots + \phi_p \gamma_{p-1} \\ &\vdots \\ \gamma_p =& \phi_1 \gamma_{p-1} + \dots + \phi_p \gamma_0 \end{aligned}$$

  which is a system of equations with p+1 equations and p+1 unknowns.

- For further autocovariances, iterate on equation (4).

Mixed Autoregressive Moving Average Processes

ARMA(p,q)

- An ARMA(p,q) process includes both autoregressive and moving average terms:

$$\begin{aligned} Y_t = & c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} \\ & + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q} \end{aligned} \tag{5}$$

or in lag operator form,

$$(1 - \phi_1 L - \cdots - \phi_p L^p) Y_t = c + (1 + \theta L + \cdots + \theta_q L^q)\epsilon$$

- Provided that the roots of:

$$1 - \phi_1 z - \cdots - \phi_p z^p = 0,$$

lie outside the unit circle, the process is stationary.

Mixed Autoregressive Moving Average Processes

ARMA(p,q)

- To find autocovariances subtract the unconditional mean ($\mu = 1/(1 - \phi_1 - \dots \phi_p)$) on both sides of equation (5) multiply by $Y_{t-j} - \mu$ and take expectations:

$$\gamma_j = \phi_1 \gamma_{j-1} + \cdots + \phi_p \gamma_{j-p} \text{ for } j > q \tag{6}$$

- For an ARMA(1,1) we have:

$$\begin{aligned} \gamma_0 &= \phi_1 \gamma_1 + \sigma^2 (1 + \theta_1^2 + \phi_1 \theta_1) \\ \gamma_1 &= \phi_1 \gamma_0 + \theta_1 \sigma^2 \\ \gamma_j &= \phi_1 \gamma_{j-1} \text{ if j} > 1 \end{aligned}$$

Mixed Autoregressive Moving Average Processes

ARMA(p,q)

- Which is a system of equations with p+1 equations and p+1 unknowns.
- For further autocovariances, iterate on equation (6).
- For estimation of ARMA models using the Kalman filter we need the first $\max\{p, q+1\}$ autocovariances.

Invertibility

- Consider an $MA(1)$ process:

$$Y_t - \mu = (1 + \theta L)\epsilon$$

- Provided the $|\theta| < 1$ we can rewrite it as a $AR(\infty)$:

$$(1 - \theta L + \theta^2 L^2 - \theta^3 L^3 + \dots)(Y_t - \mu) = \epsilon_t$$

- The process is then said invertible.

- For an $MA(q)$ the process is invertible provided that the roots of:

$$1 + \theta_1 z + \theta_2 z^2 + \dots + \theta_q z^q = 0$$

lie outside the unit circle.

Box-Jenkins Modeling Philosophy

- Box and Jenkins popularized a three-stage method aimed at selecting an appropriate model for the purpose of estimating a univariate time series:

    1. Identification: examine autocorrelation (ACF) and partial autocorrelation (PACF) function. A comparison of the samples ACF and PACF to those of various theoretical ARMA processes may suggest several plausible models.

    2. Estimation of each of the tentative models

    3. Model selection and ensure residuals mimic white-noise process.

Box-Jenkins Modeling Philosophy

Identification

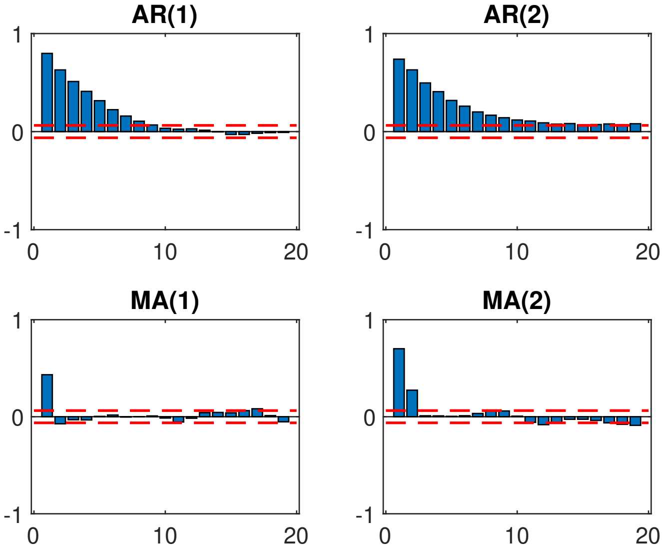
- the jth autocorrelation of a covariance-stationary process is defined as:

$$\rho_j = \frac{\gamma_j}{\gamma_0}$$

- Sample autocovariance: $\hat{\gamma}_j = \dfrac{1}{T} \sum_{j+1}^{T} (y_t - \hat{\mu})(y_{t-j} - \hat{\mu})$
- Sample autocorrelation: $\hat{\rho}_j = \dfrac{\hat{\gamma}_j}{\hat{\gamma}_0}$
- If data was generated by a white noise process: $\hat{\rho}_j \xrightarrow{d} N(0, 1/T)$

Box-Jenkins Modeling Philosophy

Identification: Autocorrelation Functions



Box-Jenkins Modeling Philosophy

Identification

- the mth partial autocorrelation is the last coefficient in an OLS regression of y on a constant and its j most recent values:

$$y_{t+1} = \hat{c} + \hat{\alpha}_1^{(m)} y_t + \hat{\alpha}_2^{(m)} y_{t-1} + \cdots + \hat{\alpha}_m^{(m)} y_{t-m+1} + \hat{e}_t$$

- If the data were really generated by a $AR(p)$ process, then the sample estimate $\hat{\alpha}_m^{(m)}$ for $m > p$ would have a variance around the true value (0) that could be approximated by:

$$Var(\hat{\alpha}_m^{(m)}) \simeq 1/T \text{ for } m > p$$

Box-Jenkins Modeling Philosophy

Identification: Partial Autocorrelation Functions