PMDS504L: Stationary Time Series Models

Autocorrelation & Partial Autocorrelation

Dr. Jisha Francis

Department of Mathematics
School of Advanced Sciences
Vellore Institute of Technology
Vellore Campus, Vellore - 632 014
India

Table of Contents

Introduction

- In statistical modeling, we aim to approximate the true relationship between inputs and outputs.
- Linear models provide a simplification to ease modeling efforts.
- A fundamental assumption is the linearity assumption .

Linear Filter

Definition: A linear filter is a linear operation that transforms one time series x_t into another time series y_t.

$$y_t = L(x_t) = \sum_{i=-\infty}^{+\infty} \psi_i x_{t-i} \tag{1}$$

- The filter acts as a process that converts an input series into an output series.
- This transformation involves past, present, and future values of the input.
- The conversion assigns different weights ψ_i to each value of x_t.

Properties of a Linear Filter

A linear filter must satisfy the following properties:

1. **Time-invariant:** The coefficients ψ_i do not depend on time.
2. **Physically realizable:** If $\psi_i = 0$ for $i < 0$, meaning the output y_t depends only on the current and past values of x_t:

$$y_t = \psi_0 x_t + \psi_1 x_{t-1} + \dots \tag{2}$$

3. **Stable:** The filter is stable if:

$$\sum_{i=-\infty}^{+\infty} |\psi_i| < \infty \tag{3}$$

Importance of Linear Filters

- Linear filters are widely used in time series analysis , particularly in signal processing and econometrics .
- They help in smoothing, forecasting, and extracting relevant features from noisy data.
- Under certain conditions, properties like stationarity of the input are preserved in the output.

Stationary Time Series Representation

For a time-invariant and stable linear filter and a stationary input time series x_t with mean $\mu_x = \mathbb{E}(x_t)$ and autocovariance function $\gamma_x(k) = \text{Cov}(x_t, x_{t+k})$, the output time series y_t is also a stationary time series with:

$$\mathbb{E}(y_t) = \mu_y = \sum_{i=-\infty}^{\infty} \psi_i \mu_x \tag{4}$$

and its autocovariance function given by:

$$\text{Cov}(y_t, y_{t+k}) = \gamma_y(k) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \psi_i \psi_j \gamma_x(i - j + k) \tag{5}$$

Stable Linear Process with White Noise

The following stable linear process with white noise ϵ_t is also stationary:

$$y_t = \mu + \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} \tag{6}$$

where $E(\epsilon_t) = 0$ and its autocovariance function is:

$$\gamma_\epsilon(h) = \begin{cases} \sigma^2, & \text{if } h = 0 \\ 0, & \text{if } h \neq 0 \end{cases} \tag{7}$$

Autocovariance Function of y_t

For y_t, the autocovariance function is:

$$\gamma_y(k) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \psi_i \psi_j \gamma_\epsilon(i - j + k) \tag{8}$$

$$= \sigma^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+k} \tag{9}$$

Proof: Autocovariance Function of y_t

Substituting y_t and y_{t+k}, we get:

$$\gamma_y(k) = E\left[\sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} \sum_{j=0}^{\infty} \psi_j \epsilon_{t+k-j}\right]. \tag{10}$$

Expanding expectation:

$$\gamma_y(k) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \psi_i \psi_j E[\epsilon_{t-i} \epsilon_{t+k-j}]. \tag{11}$$

Since ϵ_t is white noise:

$$E[\epsilon_{t-i} \epsilon_{t+k-j}] = \gamma_\epsilon(i - j + k). \tag{12}$$

Proof: Autocovariance Function of y_t (contd.)

Thus:

$$\gamma_y(k) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \psi_i \psi_j \gamma_\epsilon(i - j + k). \tag{13}$$

Since $\gamma_\epsilon(h) = \sigma^2$ if $h = 0$ and 0 otherwise, we obtain:

$$\gamma_y(k) = \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \psi_i \psi_j \sigma^2 \delta(i - j + k). \tag{14}$$

Only terms where $j = i + k$ contribute: $\gamma_y(k) = \sigma^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+k}$. This proves the final result.

Linear Process Representation

Recall: the linear process is

$$y_t = \mu + \psi_0 \epsilon_t + \psi_1 \epsilon_{t-1} + \psi_2 \epsilon_{t-2} + \dots \tag{15}$$

where:

- μ is the mean,
- ψ_i are the coefficients,
- ϵ_t is a white noise process.

Using the Backshift Operator

The backshift operator B is defined as:

$$B^i \epsilon_t = \epsilon_{t-i}$$

Rewriting the linear process using B:

$$y_t = \mu + \sum_{i=0}^{\infty} \psi_i B^i \epsilon_t \tag{16}$$

Factorizing ϵ_t

Since ϵ_t appears in each term, factor it out:

$$y_t = \mu + \left(\sum_{i=0}^{\infty} \psi_i B^i \right) \epsilon_t \tag{17}$$

Infinite Moving Average Representation

Defining:

$$\Psi(B) = \sum_{i=0}^{\infty} \psi_i B^i \tag{18}$$

we obtain the final expression:

$$y_t = \mu + \Psi(B)\epsilon_t \tag{19}$$

This is called the infinite moving average model, which represents any stationary time series.

Special Cases

The infinite moving average serves as a general class of models for any stationary time series.

This is due to a theorem by **Wold (1938)**, which states that **any** nondeterministic weakly stationary time series y_t can be represented as in Eq. (15), where the sequence $\{\psi_i\}$ satisfies:

$$\sum_{i=0}^{\infty} \psi_i^2 < \infty. \tag{20}$$

- Finite order Moving Average (MA) models: only a finite number of ψ_i are nonzero.
- Finite order Autoregressive (AR) models: ψ_i are generated using a finite number of parameters.
- ARMA models: a mixture of finite order AR and MA models.

Introduction to Moving Average (MA) Models

- Moving Average (MA) models are used for analyzing stationary time series.
- The current value y_t is a function of past white noise terms.
- An **MA(q) process** of order q is given by:

$$y_t = \mu + \epsilon_t - \theta_1 \epsilon_{t-1} - \cdots - \theta_q \epsilon_{t-q} \tag{21}$$

Properties of MA(q) Processes

- **Always stationary**, as it involves a finite sum of white noise terms.
- Suitable for modeling **short-memory processes**.
- No complex stationarity conditions

Representation Using Backward Shift Operator

- The **backward shift operator** B is defined as:

$$B^i \epsilon_t = \epsilon_{t-i} \tag{22}$$

- The MA(q) process can be rewritten as:

$$y_t = \mu + (1 - \theta_1 B - \cdots - \theta_q B^q)\epsilon_t \tag{23}$$

- Or, in compact notation:

$$y_t = \mu + \Theta(B)\epsilon_t, \quad \text{where } \Theta(B) = 1 - \sum_{i=1}^{q} \theta_i B^i \tag{24}$$

Expected Value of MA(q) Process

- Since ϵ_t is **white noise** with mean zero:

$$E(y_t) = E(\mu + \epsilon_t - \theta_1\epsilon_{t-1} - \cdots - \theta_q\epsilon_{t-q}) \tag{25}$$

- Since $E(\epsilon_t) = 0$, we get:

$$E(y_t) = \mu \tag{26}$$

Variance of the MA(q) Process

The variance is calculated as:

$$\mathrm{Var}(y_t) = \gamma_y(0) = \sigma^2 \left(1 + \theta_1^2 + \cdots + \theta_q^2\right) \tag{27}$$

Autocovariance of the MA(q) Process

The autocovariance function at lag k is:

$$\gamma_y(k) = \text{Cov}(y_t, y_{t+k}) \tag{28}$$

$$\gamma_y(k) = \begin{cases} \sigma^2(-\theta_k + \theta_1\theta_{k+1} + \cdots + \theta_{q-k}\theta_q), & k = 1, 2, \ldots, q \\ 0, & k > q \end{cases} \tag{29}$$

Autocovariance Function Definition

The autocovariance function at lag k is given by:

$$\gamma_y(k) = \text{Cov}(y_t, y_{t+k}) \tag{30}$$

Expanding y_t and y_{t+k} in terms of ε_t:

$$y_t = \varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q} \tag{31}$$

$$y_{t+k} = \varepsilon_{t+k} - \theta_1\varepsilon_{t+k-1} - \cdots - \theta_q\varepsilon_{t+k-q} \tag{32}$$

Expectation Expansion

Expanding the covariance function:

$$\gamma_y(k) = E[(\varepsilon_t - \theta_1\varepsilon_{t-1} - \cdots - \theta_q\varepsilon_{t-q}) \tag{33}$$

$$\times (\varepsilon_{t+k} - \theta_1\varepsilon_{t+k-1} - \cdots - \theta_q\varepsilon_{t+k-q})] \tag{34}$$

Using the white noise properties:

$$E[\varepsilon_t\varepsilon_{t+j}] = \begin{cases} \sigma^2, & \text{if } j = 0 \\ 0, & \text{if } j \neq 0 \end{cases} \tag{35}$$

Final Autocovariance Function

The autocovariance function is given by:

$$\gamma_y(k) = \begin{cases} \sigma^2 \left(1 + \sum_{i=1}^{q} \theta_i^2\right), & k = 0 \\ \sigma^2 \left(-\theta_k + \sum_{i=1}^{q-k} \theta_i \theta_{i+k}\right), & 1 \leq k \leq q \\ 0, & k > q \end{cases} \tag{36}$$

This completes the proof.

Autocorrelation Function of MA(q) Process

From the above equations, the autocorrelation function of the MA(q) process is given by:

$$\rho_y(k) = \frac{\gamma_y(k)}{\gamma_y(0)} = \begin{cases} \frac{-\theta_k + \theta_1\theta_{k+1} + \cdots + \theta_{q-k}\theta_q}{1 + \theta_1^2 + \cdots + \theta_q^2}, & k = 1, 2, \ldots, q \\ 0, & k > q \end{cases} \tag{37}$$

Significance of ACF for MA(q)

- The ACF "cuts off" after lag q, making it useful for identifying the MA model order.
- In real-life applications, the sample ACF $r(k)$ may not be exactly zero after lag q but should be small.

Key Takeaways of MA(q) Model

- The **MA(q) process** models a time series as a function of past white noise disturbances.
- It is **always stationary** because it involves a finite number of past disturbances.
- The **mean of the process** is simply μ.

The First-Order Moving Average Process, MA(1)

The simplest finite order Moving Average (MA) model is obtained when $q = 1$ in Eq. (21), which is given by:

$$y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} \tag{38}$$

where:

- μ is the mean of the process.
- ε_t is a white noise process with mean zero and variance σ^2.
- θ_1 is the model parameter controlling the impact of the previous error term.

Autocovariance Function of MA(1)

For the first-order moving average model, the autocovariance function is given by:

$$\gamma_y(0) = \sigma^2(1 + \theta_1^2) \tag{39}$$

$$\gamma_y(1) = -\theta_1\sigma^2 \tag{40}$$

$$\gamma_y(k) = 0, \quad k > 1 \tag{41}$$

- $\gamma_y(0)$ represents the variance of the process.
- $\gamma_y(1)$ represents the covariance at lag 1.
- For $k > 1$, the autocovariance is zero, indicating the MA(1) model does not exhibit autocorrelation beyond lag 1.

Autocorrelation Function of MA(1)

The autocorrelation function (ACF) is given by:

$$\rho_y(1) = \frac{-\theta_1}{1+\theta_1^2} \tag{42}$$

$$\rho_y(k) = 0, \quad k > 1 \tag{43}$$

- The ACF of an MA(1) process cuts off after lag 1.
- The first-lag autocorrelation, $\rho_y(1)$, depends on θ_1 and is always bounded within $[-0.5, 0]$ for $\theta_1 > 0$.
- The ACF helps in identifying the order of the moving average process.

Key Takeaways of MA(1) Model

- The MA(1) process models the dependency between consecutive observations using a single lagged error term.
- The autocovariance function confirms that the influence of past values disappears beyond lag 1.
- The autocorrelation function provides an easy way to identify MA(1) in real-world data.

The Second-Order Moving Average Process, MA(2)

Another useful finite-order moving average process is the MA(2) model, which is given by:

$$y_t = \mu + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} \tag{44}$$

Alternatively, using the backward shift operator B, we can express it as:

$$y_t = \mu + (1 - \theta_1 B - \theta_2 B^2)\varepsilon_t \tag{45}$$

Autocovariance Function of MA(2)

The autocovariance function for the MA(2) model is given by:

$$\gamma_y(0) = \sigma^2(1 + \theta_1^2 + \theta_2^2) \tag{46}$$

$$\gamma_y(1) = \sigma^2(-\theta_1 + \theta_1\theta_2) \tag{47}$$

$$\gamma_y(2) = \sigma^2(-\theta_2) \tag{48}$$

$$\gamma_y(k) = 0, \quad k > 2 \tag{49}$$

Autocorrelation Function of MA(2)

The autocorrelation function is given by:

$$\rho_y(1) = \frac{-\theta_1 + \theta_1\theta_2}{1 + \theta_1^2 + \theta_2^2} \tag{50}$$

$$\rho_y(2) = \frac{-\theta_2}{1 + \theta_1^2 + \theta_2^2} \tag{51}$$

$$\rho_y(k) = 0, \quad k > 2 \tag{52}$$

Finite Order MA Processes

- Only a few past disturbances affect the present value.
- The influence shifts over time, making older effects obsolete.
- Some processes need to consider long-term past effects.
- Estimating too many weights makes modeling difficult.

Autoregressive Models

- AR models avoid infinite weight estimation.
- They assume past values follow a pattern.
- A small set of parameters can model the series.
- Useful for efficient forecasting.

First-Order Autoregressive Process, AR(1)

- Consider the time series:

$$y_t = \mu + \sum_{i=0}^{\infty} \psi_i \epsilon_{t-i} \tag{53}$$

$$= \mu + \sum_{i=0}^{\infty} \psi_i B^i \epsilon_t \tag{54}$$

$$= \mu + \Psi(B)\epsilon_t \tag{55}$$

where $\Psi(B) = \sum_{i=0}^{\infty} \psi_i B^i$.

- Disturbances further in the past contribute less.
- Weights decrease exponentially: $\psi_i = \phi^i$.
- Condition: $|\phi| < 1$ ensures decay.

Autoregressive Models

- AR models avoid infinite weight estimation.
- They assume past values follow a pattern.
- A small set of parameters can model the series.
- Useful for efficient forecasting.

References

This presentation is adapted from:

- Montgomery, Douglas C., Cheryl L. Jennings, and Murat Kulahci. Introduction to time series analysis and forecasting. John Wiley & Sons, 2015.

Thank You!

Thank you for your attention!