BU.610.740: Forecasting Models for Business Intelligence
Chapter 3: ARIMA Models - Part I

Ali Eshragh

Spring II, 2024

Outline

Core Concepts for Time Series Analysis

- Chapter 0 presented the main concepts of the **correlation coefficient** and the properties of the **normal distribution**, essential for **linear** statistical modeling.
- Chapter 1 introduced the fundamental concept of **stationarity**, along with two important tools: **autocorrelation** and **cross-correlation** functions, to examine significant **linear** relationships in time series across **time lags**.
- Chapter 2 discussed **linear regression** modeling in **time series** contexts and **recent advances** in addressing challenges with **big data** problems.

Chapter 3

This chapter **applies** all these concepts and models to **develop** statistical regression techniques for **time series** data analysis.

Autoregressive Model

Definition (Autoregressive Model)

An **autoregressive** model of **order** p, denoted by $\mathtt{AR}(p)$, is described by

$$X_t = \phi_0 + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + W_t$$

where X_t is the value of a **stationary** time series at time t, $\phi_0, \phi_1, \ldots, \phi_p$ are the model's **coefficients** with $\phi_p \neq 0$ ensuring the model's order, and W_t represents **Gaussian white noise** with zero **mean** and constant **variance** σ_W^2.

AR Model: Exploring Time Series Dynamics

- An AR(p) model **suggests** that the time series **current** value X_t can be modeled by a **linear** combination of the p **lagged** values $X_{t-1}, X_{t-2}, \ldots, X_{t-p}$
- The model highlights the **auto** in **autoregression**, where it implies **self** in **regression**.
- An AR(p) model can **only** be applied to **stationary** time series data, as it may **not** effectively capture variations in a **non-stationary** time series.

$\mathtt{AR}(1)$ Model: Case $|\phi_1| < 1$

- **Consider** the $\mathtt{AR}(1)$ model $X_t = \phi_0 + \phi_1 X_{t-1} + W_t$ for $|\phi_1| < 1$
- It can be **shown** that X_t can be represented by the **current** and **previous** values of the **white noise** series as

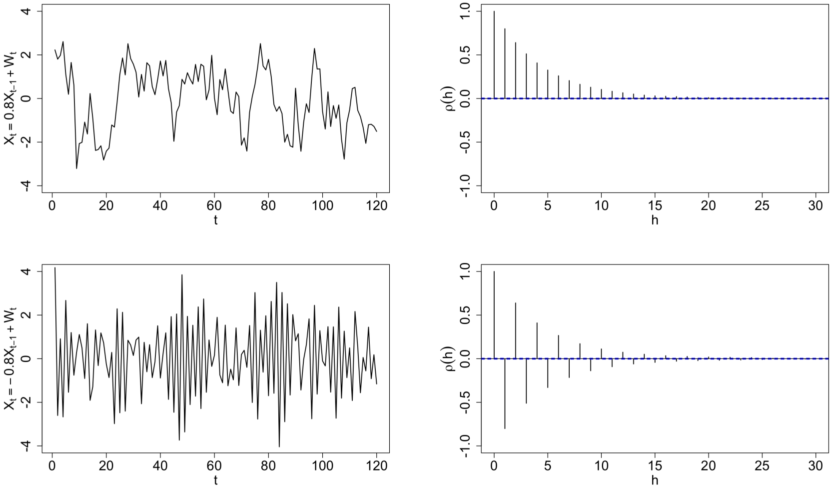
$$\begin{aligned} X_t &= \frac{\phi_0}{1-\phi_1} + \sum_{j=0}^{\infty} \phi_1^j W_{t-j} \\ &= \frac{\phi_0}{1-\phi_1} + W_t + \phi_1 W_{t-1} + \phi_1^2 W_{t-2} + \cdots \end{aligned}$$

- The **mean** function equals $\mathbb{E}[X_t] = \dfrac{\phi_0}{1-\phi_1}$
- The **autocorrelation** function at lag h is given by

$$\rho_X(h) = \phi_1^h \quad \text{for } h = 0, 1, 2, \ldots$$

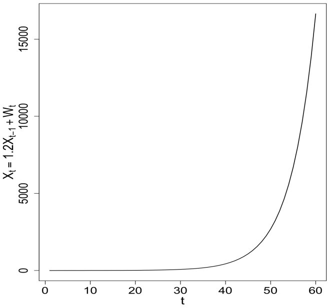
- X_t is **correlated** with all **lagged** values of the time series.

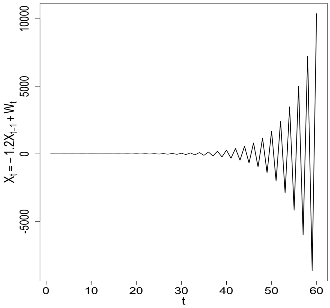
Visualization of AR(1) Model Properties



$\texttt{AR}(1)$ Model: Case $|\phi_1| \geq 1$

- **Consider** the $\texttt{AR}(1)$ model $X_t = \phi_0 + \phi_1 X_{t-1} + W_t$ for $|\phi_1| \geq 1$
- Due to $|\phi_1| \geq 1$, in **contrast** to when $|\phi_1| < 1$, X_t **cannot** be represented by the **current** and **previous** values of the **white noise** series.
- This process is called an **explosive** process, as the **value** of the time series rapidly **grows** in magnitude.





Causal Autoregressive Models

Definition (Causal Models)

An $\texttt{AR}(p)$ model $X_t = \phi_0 + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + W_t$ is **causal** if X_t can be expressed as a **linear combination** of the **current** and **previous** values of the **white noise** series as

$$
\begin{aligned}
X_t &= \mu_X + W_t + \sum_{j=1}^{\infty} c_j W_{t-j} \\
&= \mu_X + W_t + c_1 W_{t-1} + c_2 W_{t-2} + \cdots
\end{aligned}
$$

where μ_X is the **mean function** of the **stationary** time series, $\mathbb{E}[X_t]$, and c_i are **constant coefficients**, determining the **influence** of **past** white noise values on the **current** value of the time series.

Causal Models: Stability and Predictability

- If an autoregressive model is **causal**, then it is **stable** and **predictable**, like an $\mathtt{AR}(1)$ model with $|\phi_1| < 1$

- Conversely, if an autoregressive model is **not** causal, then it is potentially **unstable** or **explosive**, making its future **predictions** increasingly **off-target** and **less reliable**.

- An $\mathtt{AR}(p)$ model $X_t = \phi_0 + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + W_t$ is **causal** if and only if **all roots** of the polynomial equation

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

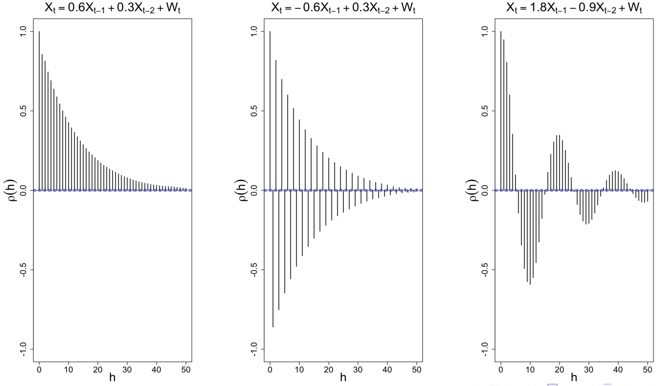
have **absolute** values greater than one.

- Henceforth, in this course, we **assume** that all **autoregressive** models are **causal**, unless **otherwise** stated.

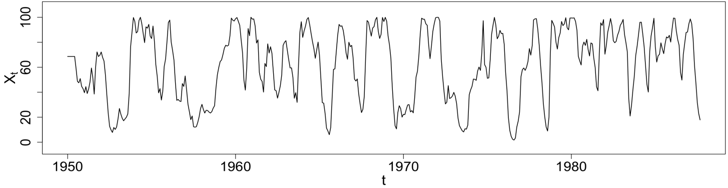
Autocorrelation Patterns in Causal AR Models

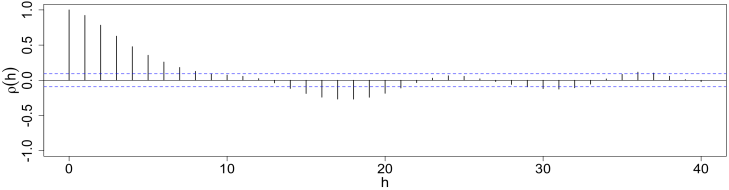
- The **autocorrelation** function of a **causal** $\texttt{AR}(p)$ model exhibits either **exponential decay** or a **sinusoidal pattern** as the **lag** increases.

- X_t is **correlated** with all **lagged** values of the time series, like an $\texttt{AR}(1)$ model with $|\phi_1| < 1$



Case Study: Recruitment Time Series Data





Moving Average Model

Definition (Moving Average Model)

A **moving average** model of **order** q, denoted by $\mathtt{MA}(q)$, is described by

$$X_t = \theta_0 + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q} + W_t$$

where X_t is the value of a **stationary** time series at time t, $\theta_0, \theta_1, \ldots, \theta_q$ are the model's **coefficients** with $\theta_q \neq 0$ ensuring the model's order, and W_t represents **Gaussian white noise** with zero **mean** and constant **variance** σ_W^2.

MA Model: Exploring Time Series Dynamics

- An $\mathtt{MA}(q)$ model **suggests** that the time series **current** value X_t can be modeled by a **linear** combination of the q **lagged** values of **white noise** terms $W_{t-1}, W_{t-2}, \ldots, W_{t-q}$
- The model highlights the **moving average** component, suggesting that the **current** value of the time series is determined by a **weighted sum** of past white noise terms calculated over a **moving window**.
- An $\mathtt{MA}(q)$ model can **only** be applied to **stationary** time series data, as it may **not** effectively capture variations in a **non-stationary** time series.

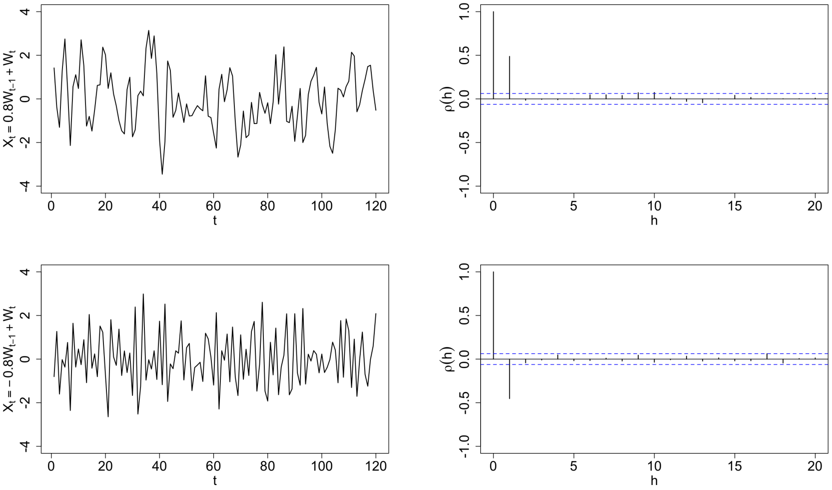
MA(1) Model: Features and Characteristics

- **Consider** the MA(1) model $X_t = \theta_0 + \theta_1 W_{t-1} + W_t$
- The **mean** function equals $\mathbb{E}[X_t] = \theta_0$
- The **autocorrelation** function at lag h is given by

$$\rho_X(h) = \begin{cases} \dfrac{\theta_1}{1+\theta_1^2} & \text{for } h = 1 \\ 0 & \text{for } h = 2, 3, \ldots \end{cases}$$

- X_t is **correlated** with X_{t-1}, but **not** with $X_{t-2}, X_{t-3}, \ldots$
- This **contrasts** with the AR(1) model, where the **correlation** between X_t and X_{t-h} is **never** zero for $h = 1, 2, \ldots$

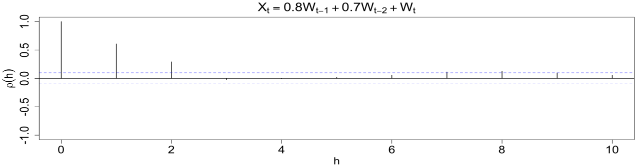
Visualization of MA(1) Model Properties

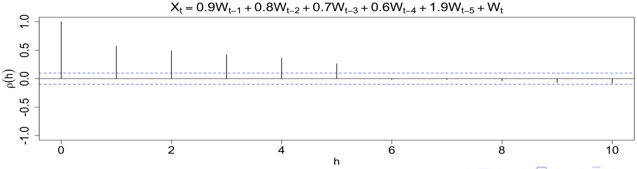


Autocorrelation Patterns in MA Models

- The **autocorrelation** function of an $\mathtt{MA}(q)$ model **resembles** that of an $\mathtt{MA}(1)$ model, with non-zero values at **lag** q and possibly **earlier** lags, but zeros from lag $q+1$ **onwards**.

- X_t is **correlated** with X_{t-q} and possibly **earlier** lags, but **not** with $X_{t-(q+1)}, X_{t-(q+2)}, \ldots$





Order Estimation: MA vs. AR Models

- When the time series follows a **moving average** process, the ACF provides considerable **information** about the **order** of the model.
- The last non-zero **lag** at which the ACF **cuts off** can serve as a good initial **estimate** for the **order** of the moving average model.
- However, if the time series follows an **autoregressive** model, the ACF alone may **not** provide **sufficient** information to estimate the **order** of the model.
- Consequently, we may need to **introduce** a new function that **behaves** similarly to the `ACF` of **moving average** models, but is **tailored** for **autoregressive** models.

Partial Autocorrelation Function

- The **partial autocorrelation function** (PACF) is designed to **address** the lack of a function to **estimate** the order of **autoregressive** models.

- It is **defined** over lags $1, 2, \ldots$, and, similar to the ACF, it **takes** values within $[-1, 1]$

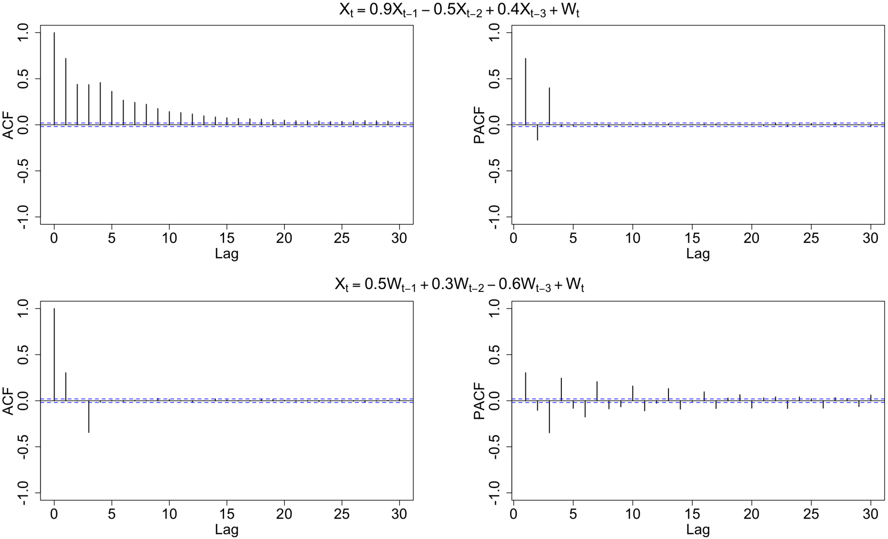
Properties

- For a **causal** $\mathtt{AR}(p)$ model $X_t = \phi_0 + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + W_t$, the PACF at lag p **equals** $\phi_p (\neq 0)$ and may **exhibit** non-zero values at earlier lags $1, \ldots, p-1$, but **drops** to zeros from lag $p+1$ onwards.

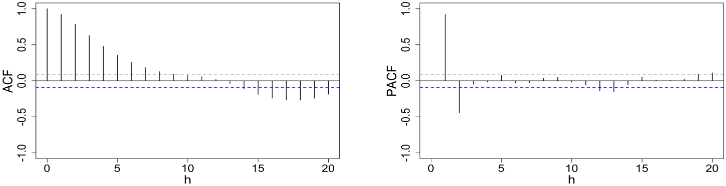
- For an $\mathtt{MA}(q)$ model, the PACF is non-zero at **all** lags and **approaches** zero as the number of lags **tends** to infinity.

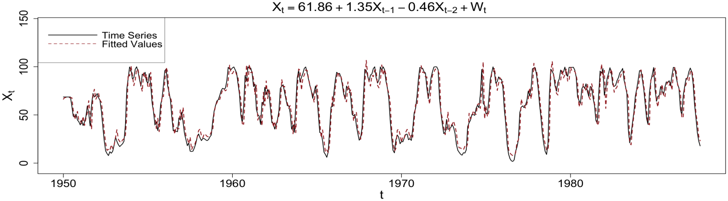
- Therefore, the PACF in AR models is **similar** to the ACF in MA models, and the ACF in AR models **resembles** the PACF in MA models.

Visualizing ACF and PACF in AR and MA Models



Case Study: Recruitment Time Series Data





Autoregressive Moving Average Model

Definition (Autoregressive Moving Average Model)

An **autoregressive moving average** model of **orders** p and q, denoted by $\texttt{ARMA}(p, q)$, is described by

$$X_t = \phi_0 + \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q} + W_t$$

where X_t is the value of a **stationary** time series at time t, $\phi_p \neq 0$ and $\theta_q \neq 0$ **ensuring** the model's orders, and W_t represents **Gaussian white noise** with zero **mean** and constant **variance** σ_W^2.

ARMA Model Fundamentals

- An $\texttt{ARMA}(p, q)$ model **comprises** two components, the $\texttt{AR}(p)$ part, using p lagged values of the **time series**, and the $\texttt{MA}(q)$ part, involving q lagged **white noise** values, together **modeling** the current series value X_t as their **linear combination**.

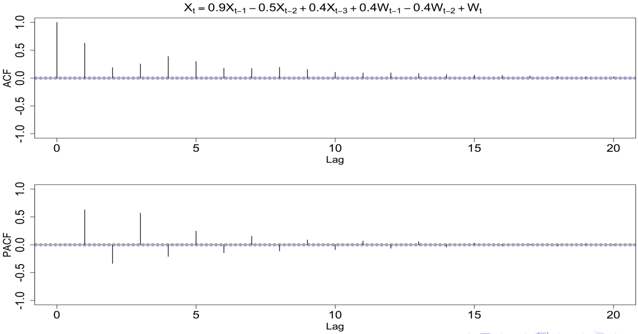
- An $\texttt{ARMA}(0, q)$ model with $p = 0$ **simplifies** to a mere $\texttt{MA}(q)$ model.

- An $\texttt{ARMA}(p, 0)$ model with $q = 0$ **simplifies** to a mere $\texttt{AR}(p)$ model.

- An $\texttt{ARMA}(p, q)$ model can **only** be applied to **stationary** time series data, as it may **not** effectively capture variations in a **non-stationary** time series.

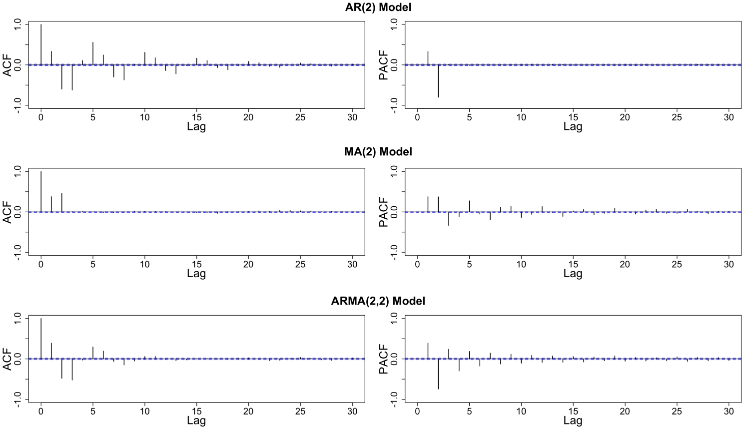
ACF and PACF Patterns in ARMA Models

- The ACF and PACF **behaviors** in an ARMA model reflect a **mix** of AR and MA influences, showing characteristics of **both** models in one:
  - The ACF exhibits a **gradual decline**, influenced by both AR and MA components, **without** a sharp cutoff.
  - The PACF presents a **complex pattern** due to the interplay of AR and MA effects, **lacking** distinct cutoffs.



ACF and PACF Across AR, MA, and 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**



Assignment 4

Question 1

Consider an $\texttt{MA}(\infty)$ model

$$\begin{aligned} X_t &= \theta_0 + W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \cdots \\ &= \theta_0 + W_t + \sum_{i=1}^{\infty} \theta_i W_{t-i} \end{aligned}$$

Show that the **causal** $\texttt{AR}(1)$ model $X_t = \phi_0 + \phi_1 X_{t-1} + W_t$ for $|\phi_1| < 1$ can be **represented** as an $\texttt{MA}(\infty)$ model. What are the **coefficients** θ_i in terms of ϕ_0 and ϕ_1 in that model?

Understanding ACF and PACF Patterns in AR, MA, and ARMA Models

- A **causal** $\mathtt{AR}(p)$ model can be **represented** as an $\mathtt{MA}(\infty)$ model.
- This explains why the ACF for an AR model displays a **tails off** pattern, as it effectively has an infinite MA **order**, leading to **no sharp** cutoff.
- Under certain mild **conditions**, it can be demonstrated that an $\mathtt{MA}(q)$ model can be **represented** as an $\mathtt{AR}(\infty)$ model.
- This explains why the PACF for an MA model displays a **tails off** pattern, as it effectively has an infinite AR **order**, leading to **no sharp** cutoff.
- This insight **clarifies** why the ACF and PACF of an $\mathtt{ARMA}(\mathtt{p}, \mathtt{q})$ model, which **comprises** both $\mathtt{AR}(p)$ and $\mathtt{MA}(q)$ components, exhibit a **complex pattern** that is a **blend** of the two.

References

1. R.H. Shumway and D.S. Stoffer, *Time Series Analysis and Its Applications With R Examples*, Springer, New York, 2017.