Time Series Analysis and Modelling

Part 5: Time Series Models (ARIMA)

Jonathan Mwaura

Khoury College of Computer Sciences

August 1, 2024

The Roux Institute
at Northeastern University

Introduction

Acknowledgements

- These slides have been adapted from Achraf Cohen, University of West Florida - from the class class STA6856
- Lecture Notes from Dewei Wang, Department of Statistics, University of South Carolina (See notes in Canvas)
- C. Chatfield, The Analysis of Time Series: Theory and Practice, Chapman and Hall (1975).

ARIMA models

We have already seen the importance of the class of ARMA models for representing stationary time series. A generalization of this class, which includes a wide range of nonstationary series, is provided by the ARIMA (AutoRegressive Integrated Moving Average)models.

Definition

If d is a nonnegative integer, then X_t is an **ARIMA(p,d,q) process** if

$$Y_t := (1 - B)^d X_t$$

is a causal ARMA(p,q) process.
B is the backward shift operator.

- *ARIMA processes reduce to ARMA processes when differenced finitely many times*
- X_t *is stationary if and only if* $d = 0$

ARIMA models

The definition means that X_t satisfies a difference equation of the form:

$$\phi(B)(1-B)^d X_t = \theta(B) Z_t, \quad Z_t \sim WN(0, \sigma^2)$$

The polynomial $\phi^*(B) = \phi(B)(1-B)^d$ has now a zero of order d at $z = 1$.

Example

Consider X_t is an ARIMA(1,1,0) process of for $\phi \in (-1, 1)$,

$$(1-\phi B)(1-B)X_t = Z_t, \quad Z_t \sim WN(0, \sigma^2)$$

ARIMA models

ACF for ARIMA models

- A distinctive feature of the data that suggests the appropriateness of an ARIMA model is **the slowly decaying positive sample autocorrelation function**.
- In order to fit an ARIMA model to the data, we would apply the operator $\nabla = 1 - B$ repeatedly in the hope of some j, $\nabla^j X_t$ will have a **rapidly decaying sample autocorrelation function**, that is compatible with that of an ARMA process with no zeros of the autoregressive polynomial near the unit circle.

Example with R with $\phi = 0.8$, n=200, and $\sigma^2 = 1$

ARIMA models

Modeling of ARIMA models

- Deviations (e.g. trend, seasonality, heteroskedasticity) from stationary may be suggested by the graph of the the series itself or by the sample autocorrelation function or both.
- We have seen how to handle the Trend and Seasonality components
- Logarithmic Transformation is appropriate whenever the series whose variance increases linearly with the mean. A general class of variance-stabilizing transformations is given by **Cox-Box transformation** f_λ:

$$f_\lambda(X_t) = \begin{cases} \lambda^{-1}(X_t^\lambda - 1), & X_t \geq 0, \lambda > 0 \\ \ln X_t, & X_t > 0, \lambda = 0 \end{cases}$$

- In practice, λ is often 0 or 0.5
-

ARIMA models

Units Roots in Time Series

- The unit root problem arises when either the AR or MA polynomial of ARMA model has a root on or near to the unit circle.
- *A root near to 1 of the AR* polynomial suggest that the data **should be differenced** before fitting an ARMA model
- *A root near to 1 of the MA* polynomial suggest that the data **were overdifferenced.**

ARIMA models

Units Roots in Time Series:

- Augmented Dickey-Fuller Test for AR processes
- For MA process is more complicated (general case not fully resolved)

SARIMA models

We have already seen the how difference the series X_t at lag s is convenient way of eliminating a seasonal component of period s.

Definition

If d and D are nonnegative integers, then X_t is a seasonal $ARIMA(p, d, q)x(P, D, Q)_s$ process with period s if the difference series $Y_t = (1 - B)^d (1 - B^s)^D X_t$ is a causal ARMA process defined by

$$\phi(B)\Phi(B^s)Y_t = \theta(B)\Theta(B^s)Z_t$$

where B is the backward shift operator. $Z_t \sim WN(0, \sigma^2)$.
$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$, $\Phi(z) = 1 - \Phi_1 z - \cdots - \Phi_P z^P$,
$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$, and $\Theta(z) = 1 + \Theta_1 z + \cdots + \Theta_Q z^Q$

- *The process Y_t is causal if and only if $\phi(z) = 0$ and $\Phi(z) = 0$ for $\mid z \mid > 1$*

SARIMA models

A nonstationary process has often a seasonal component that repeats itself after a regular period of time. The seasonal period can be:

- Monthly: s=12 (12 observations per a year)
- Quarterly: s=4 (4 observations per a year)
- Daily: s=365 (365 observations per a year)
- Daily per week: s=5 (5 working days)
- Weekly: s=52 (observations per a year)