Time Series Analysis and Modelling

Part 4: Time Series Models (ARMA)

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).

ARMA(p,q) models

$ARMA(p, q)$ process

X_t is an ARMA(p, q) process if X_t is stationary and if for every t,

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}$$

where $Z_t \sim WN(0, \sigma^2)$ and the polynomials $(1 - \phi_1 z - \cdots - \phi_p z^p)$ and $(1 + \theta_1 z + \cdots + \theta_q z^q)$ have no common factors.

The process X_t is said to be an ARMA(p,q) process with mean μ if $X_t - \mu$ is an ARMA(p,q) process.

ARMA Models

We can write:

$$\phi(B)X_t = \theta(B)Z_t$$

where $\phi(.)$ and $\theta(.)$ are the *pth* and *qth* degree polynomials:

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

and

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$$

and B is the backward shift operator $B^j X_t = X_{t-j}$, $j = 0, \pm 1, ...$

ARMA models

$ARMA(1,1)$ process

X_t is an ARMA$(1,1)$ process if X_t is stationary and if for every t,

$$X_t - \phi_1 X_{t-1} = Z_t + \theta_1 Z_{t-1}$$

where $Z_t \sim WN(0, \sigma^2)$

ARMA models

$ARMA(p, q)$ process: Existence, Causality, and Invertibility

- A stationary solution (existence, uniqueness) X_t exists if and only if:

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0 \quad \text{for all} \mid z \mid = 1$$

- An ARMA(p,q) process X_t is **causal** if there exist constant ψ_j such that $\sum_{j=0}^{\infty} \mid \psi_j \mid < \infty$ and $X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$, for all t. Causality is equivalent to the condition:

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p \neq 0 \quad \text{for all} \mid z \mid \leq 1$$

- An ARMA(p,q) process X_t is **invertible** if there exist constant π_j such that $\sum_{j=0}^{\infty} \mid \pi_j \mid < \infty$ and $Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}$, for all t. Invertibility is equivalent to the condition:

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q \neq 0 \text{ for all} \mid z \mid \leq 1.$$

ARMA models

$ARMA(p, q)$ process: Autocorrelation Function (ACF)

The autocorrelation function (ACF) of the causal ARMA(p,q) process X_t can be found using the fact that $X_t = \sum_{j=0}^{\infty} \psi_j Z_{t-j}$ (causal MA(∞) process) and

$$\gamma_X(h) = \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+|h|}$$

Find the ACF of ARMA(1,1), using the fact that

$$\psi_0 = 1; \psi_j = (\phi + \theta)\phi^{j-1}, \quad j \geq 1$$

ARMA models

$ARMA(1,1)$ process: Autocorrelation Function (ACF)

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

$$\gamma_X(1) = \frac{(\theta+\phi)(1+\phi\theta)}{1-\phi^2}\sigma^2$$

$$\gamma_X(h) = \phi^{h-1}\gamma_x(1)$$

ARMA models

Definition: Partial Autocorrelation Function (PACF)

The partial autocorrelation function (PACF) of ARMA process X_t is the function $\alpha(.)$ defined by

$$\alpha(0) = 1$$

and

$$\alpha(h) = \phi_{hh}, \quad h \geq 1$$

where ϕ_{hh} is the last component of $\boldsymbol{\phi_h} = \Gamma_h^{-1}\boldsymbol{\gamma}_h$

Think about it as a conditional correlation $Cor(X_t, X_{t+h} \mid X_{t+1}, \dots, X_{t+h-1})$

Modeling ARMA models

Practical facts about PACF and ACF

- The PACF is often best used to identify the AR(p) models
- For AR(p) models, the theoretical PACF are equal to zero after $h = p$
- The MA(q) models are better identified using ACF.

Modeling ARMA models

To determine an appropriate ARMA(p,q) model to represent an observed stationary process, we need to:

- Choose the orders p and q (order selection)
- Estimate the mean
- Estimate the coefficients $\{\phi_i, i = 1, \ldots, p\}$ and $\{\theta_i, i = 1, \ldots, q\}$
- Estimate the white noise variance σ^2
- Select a model

Modeling ARMA models

To determine an appropriate ARMA(p,q) model to represent an observed stationary process, we need to:

- Choose the orders p and q (order selection)Use ACF and PACF plots.
- Estimate the mean Use the mean-corrected process $X_t - \overline{X}_n$.
- Estimate the coefficients $\{\phi_i, i = 1, \ldots, p\}$ and $\{\theta_i, i = 1, \ldots, q\}$
- Estimate the white noise variance σ^2
- Select a model

Modeling ARMA models

When p and q are **known** and the time series is **mean-corrected**, good estimators of vectors $\boldsymbol{\phi}$ and $\boldsymbol{\theta}$ can be found imaging the data to a stationary Gaussian time series and **maximizing the likelihood** with respect to the $p+q+1$ parameters ($\phi_1, \ldots, \phi_p, \theta_1, \ldots, \theta_q$ and σ^2). We can estimate these parameters using:

- The Yule-Walker and Burg procedures for pure autoregressive models AR(p)
- The Innovations and Hannan-Rissanen procedures for ARMA(p,q)

Modeling ARMA models

Properties

- The Burg's algorithm usually gives higher likelihoods than the Yule-Walker equations for AR(p)
- For pure MA processes, the Innovations algorithm usually gives higher likelihoods than the Hannan-Rissanen procedure
- For ARMA models, the Hannan-Rissanen is more successful in finding causal models

These preliminary estimations are required for initialization of the likelihood maximization.

Modeling ARMA models

Yule-Walker Estimation

The Sample Yule-Walker equations are:

$$\hat{\phi} = (\hat{\phi}_1, \ldots, \hat{\phi}_p)' = \hat{R}_p^{-1}\hat{\rho}_p$$

and

$$\hat{\sigma}^2 = \hat{\gamma}(0)\left[1 - \hat{\rho}_p'\hat{R}_p^{-1}\hat{\rho}_p\right]$$

where $\hat{\rho}_p = (\hat{\rho}(1), \ldots, \hat{\rho}(p))'$

- Large-Sample Distribution of $\hat{\phi} \sim N(\phi, \frac{\sigma^2\Gamma_p^{-1}}{n})$
- Confidence intervals of ϕ_{pj} are $\hat{\phi}_{pj} \pm z_{1-\alpha/2}\sqrt{\frac{\hat{\sigma}^2\hat{\Gamma}_p^{-1}}{n}}$

Estimation ARMA models

We can use the Innovations algorithm in order to estimate parameters of MA(q). The Confidence regions of the coefficients:

MA(q) Estimation

- Confidence intervals of θ_j are $\hat{\theta}_{mj} \pm z_{1-\alpha/2}\sqrt{\frac{\sum_{i=0}^{j-1} \hat{\theta}_{mi}^2}{n}}$

Estimation ARMA Models

Suppose X_t is a Gaussian time series with mean zero

Maximum Likelihood Estimators

$$\hat{\sigma}^2 = \frac{S(\hat{\phi}, \hat{\theta})}{n}$$

where $S(\hat{\phi}, \hat{\theta}) = \sum_{j=1}^{n} \frac{(X_j - \hat{X}_j)^2}{r_{j-1}}$,
and $\hat{\phi}, \hat{\theta}$ are the values of ϕ, θ that minimize:

$$l(\phi, \theta) = \ln n^{-1} S(\phi, \theta) + n^{-1} \sum_{j=1}^{n} \ln r_{j-1}$$

Minimization of $l(\phi, \theta)$ must be done numerically. Initial values for $\phi, \theta)$ can be obtained from preliminary estimation algorithms (Yule-walker, Burg, Innovations, and Hannan).

ARMA Models

Order selection

The Akaike Information criterion bias-corrected (AICC) is defined as follows:

$$AICC := -2\ln(L_X(\beta, S_X(\beta)/n) + \frac{2(p+q+1)n}{n-p-q-2}$$

It was designed to be an approximately unbiased estimate of the Kullback-Leibler index of the fitted model relative to the true model.

We select p and q values for our fitted model to be those that minimize $AICC(\hat{\beta})$.

Modeling ARMA models

Paramater Redundancy

Consider a white noise process $X_t = Z_t$. We can write this as

$$0.3X_{t-1} = 0.3Z_{t-1}$$

By subtracting the two representation we have:

$$X_t - 0.3X_{t-1} = Z_t - 0.3Z_{t-1}$$

which looks like an ARMA(1,1) model. Of course X_t is still white noise. We have this problem because of the parameter redundancy or over-parameterization.
We can solve this problem by looking at the common factors of the two polynomials $\phi(B)$ and $\theta(B)$.

Forecasting ARMA models

Given $X_1, X_2, \ldots, X_n$ observations, we want to predict X_{n+h}. We know that the best linear predictor is given by:

$$P_n X_{n+h} = \hat{X}_{n+h} = \mu + \sum_{i=1}^{n} a_i (X_{n+i-1} - \mu)$$

where the vector a_n satisfies $\Gamma_n a_n = \gamma_n(h)$. The mean squared errors:

$$E(X_{n+h} - P_n X_{n+h})^2 = \gamma(0) - a_n' \gamma(h) = \gamma(0)(1 - a_n' \rho(h))$$

Use Durbin-Levinson and Innovations algorithms to solve these equations.

Forecasting ARMA models

Examples

- For AR(1): $P_n X_{n+1} = \phi X_n$
- For AR(1) with nonzero-mean: $P_n X_{n+h} = \mu + \phi^h (X_n - \mu)$
- For AR(p): if n>p then $P_n X_{n+1} = \phi_1 X_n + \cdots + \phi_p X_{n+1-p}$
- For MA(q): $P_n X_{n+1} = \sum_{j=1}^{\min(n,q)} \theta_{nj} (X_{n+1-j} - \hat{X}_{n+1-j})$
- For ARMA(p,q): if $n > m = \max(p, q)$ then for all $h \geq 1$

$$P_n X_{n+h} = \sum_{i=1}^{p} \phi_i P_n X_{n+h-i} + \sum_{j=h}^{q} \theta_{n+h-1,j} (X_{n+h-j} - \hat{X}_{n+h-j})$$