Lecture 8 ARIMA Models

Shiwei Lan[1]

[1]School of Mathematical and Statistical Sciences
Arizona State University

STP598 Spatiotemporal Analysis
Fall 2020

Introduction of correlation

- Recall that we write time series x_t in the simple additive format

$$x_t = s_t + v_t \tag{1}$$

  where s_t denotes some unknown signal and v_t denotes a time series that may be white or correlated over time.

- In the *trend stationary* model, the process has stationary behavior around a trend:

$$x_t = \mu_t + y_t \tag{2}$$

  where x_t are the observations, μ_t denotes the trend, and y_t is a stationary process.

- We could model trend μ_t using a linear model $\mu_t = \beta_0 + \beta_1 t$.

Introduction of correlation

- Classical regression models, developed for the static case, only allow the dependent variable to be influenced by current values of the independent variables, which is insufficient.
- In the time series case, it is desirable to allow the dependent variable to be influenced by the past values of the independent variables and possibly by its own past values.
- The introduction of correlation may be generated through lagged linear relations.
- This leads to proposing the *autoregressive (AR)* and *autoregressive moving average (ARMA)* models (Whittle 1951).
- Adding nonstationary models to the mix leads to the *autoregressive integrated moving average (ARIMA)* model (Box and Jenkins 1970).

Table of Contents

Autoregressive Models

- Autoregressive models are based on the idea that the current value of the series, x_t, can be explained as a function of p past values, $x_{t-1}, x_{t-2}, \cdots, x_{t-p}$.
- An **autoregressive model** of order p, denoted as **AR**(p), is of the form

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + w_t \tag{3}$$

  where x_t is stationary, $w_t \sim wn(0, \sigma_w^2)$, and $\phi_1, \cdots, \phi_p$ are constants ($\phi_p \neq 0$).
- If the mean, μ, of x_t is not zero, we replace x_t by $x_t - \mu$ and write

$$x_t = \alpha + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + w_t \tag{4}$$

  where $\alpha = \mu(1 - \phi_1 - \cdots - \phi_p)$.
- Introducing the **autoregressive operator**, we write

$$\phi(B)x_t = w_t, \quad \phi(B) = 1 - \sum_{i=1}^{p} \phi_i B^i \tag{5}$$

Autoregressive Models

- Consider the AR(1) model

$$x_t = \phi x_{t-1} + w_t \tag{6}$$

- we could use backward substitution to get

$$x_t = \phi x_{t-1} + w_t = \phi(\phi x_{t-2} + w_{t-1}) + w_t = \cdots = \phi^k x_{t-k} + \sum_{j=0}^{k-1} \phi^j w_{t-j} \tag{7}$$

- Assuming $|\phi| < 1$ and $\sup_t \mathrm{Var}(x_t) < \infty$, we get the following linear process

$$x_t = \sum_{j=0}^{\infty} \phi^j w_{t-j} \tag{8}$$

- What is the autocovariance? Autocorrelation function (ACF)?

Autoregressive Models

- First $\mathrm{E}(x_t) = \sum_{j=0}^{\infty} \phi^j \mathrm{E}(w_{t-j}) = 0$. Second, the autocovariance

$$\gamma(h) = \mathrm{Cov}(x_{t+h}, x_t) = \sigma_w^2 \sum_{j=0}^{\infty} \phi^{h+j} \phi^j = \frac{\sigma_w^2 \phi^h}{1-\phi^2}, \quad h \geq 0 \tag{9}$$

- Then the ACF of an AR(1) is

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \phi^h, \quad h \geq 0 \tag{10}$$

- Note that $\rho(h)$ satisfies the recursion

$$\rho(h) = \phi \rho(h-1), \quad h = 1, 2, \cdots \tag{11}$$

Autoregressive Models

- Now if $\phi = 1$, is $x_t = x_{t-1} + w_t$ stationary?
- What if $|\phi| > 1$? Such processes are called explosive because the values of the time series quickly become large in magnitude.
- However, using the forward substitution we get

$$\begin{aligned} x_t &= \phi^{-1}x_{t+1} - \phi^{-1}w_{t+1} = \phi^{-1}(\phi^{-1}x_{t+2} - \phi^{-1}w_{t+2}) - \phi^{-1}w_{t+1} \\ &= \cdots = \phi^{-k}x_{t+k} + \sum_{j=1}^{k-1} \phi^{-j} w_{t+j} \end{aligned}$$

- Under the same assumption, we have the process in terms of its future

$$x_t = -\sum_{j=1}^{\infty} \phi^{-j} w_{t+j} \tag{12}$$

- When a process does not depend on the future, we say it is *causal*.

Autoregressive Models

Example

For the non-causal stationary process

$$x_t = \phi x_{t-1} + w_t, \quad |\phi| > 1 \tag{13}$$

and $w_t \overset{iid}{\sim} N(0, \sigma_w^2)$. *What is the autocovariance? ACF?*

Autoregressive Models

- To express AR(1) in linear process, we could also consider matching coefficients

$$\phi(B)x_t = w_t, \quad \phi(B) = 1 - \phi B \tag{14}$$

- We could write

$$x_t = \psi(B)w_t = \sum_{j=0}^{\infty} \psi_j w_{t-j} \tag{15}$$

- Then we have $\phi(B)\psi(B) = 1$, which implies

$$\psi_1 = \phi, \quad \psi_j = \psi_{j-1}\phi \tag{16}$$

  and it yields $\psi_j = \phi^j$.

- Another way to obtain this result is by the following series expansion

$$\phi^{-1}(z) = \frac{1}{1 - \phi z} = \sum_{j=0}^{\infty} \phi^j z^j, \quad |z| \leq 1 \tag{17}$$

Moving Average Models

- Alternative to the autoregressive representation, x_t can be a linear combination of *white noise* $\{w_t\}$.
- The **moving average model** of order q, or **MA**(q), is defined

$$x_t = w_t + \theta_1 w_{t-1} + \cdots + \theta_q w_{t-q} \tag{18}$$

where $w_t \sim wn(0, \sigma_w^2)$ and $\theta_1, \cdots, \theta_q (\theta_q \neq 0)$ are paremeters

- Introducing the **moving average operator**, we can write

$$x_t = \theta(B) w_t, \quad \theta(B) = \sum_{i=0}^{q} \theta_i B^i \tag{19}$$

where B is the backward operator such that $B^i w_t = w_{t-i}$.

Moving Average Models

- Consider the MA(1) model

$$x_t = w_t + \theta w_{t-1} \tag{20}$$

- Then $\mathrm{E}(x_t) = 0$, and the autocovariance

$$\gamma(h) = \begin{cases} (1+\theta^2)\sigma_w^2, & h = 0 \\ \theta\sigma_w^2, & h = 1 \\ 0, & h > 1 \end{cases} \tag{21}$$

and the ACF is

$$\rho(h) = \begin{cases} \frac{\theta}{1+\theta^2}, & h = 1 \\ 0, & h > 1 \end{cases} \tag{22}$$

- But how do we distinguish between

$$x_t = w_t + \frac{1}{5} w_{t-1},\ w_t \overset{iid}{\sim} N(0, 25) \quad vs. \quad y_t = v_t + 5 v_{t-1},\ v_t \overset{iid}{\sim} N(0, 1)? \tag{23}$$

Moving Average Models

- We will choose the model with an infinite AR representation. Such a process is called an *invertible* process.
- We reverse the roles of x_t and w_t:

$$w_t = -\theta w_{t-1} + x_t \tag{24}$$

which has an infinite AR representation when $|\theta| < 1$:

$$w_t = \sum_{j=0}^{\infty} (-\theta)^j x_{t-j} \tag{25}$$

- In general, we write MA process as $w_t = \pi(B)x_t$, where $\pi(B) = \theta^{-1}(B)$.
- For MA(1), if $|\theta| < 1$, we have

$$\pi(B) = \theta^{-1}(B) = (1 + \theta B)^{-1} = \sum_{j=0}^{\infty} (-\theta)^j B^j \tag{26}$$

Autoregressive Moving Average Models

- A time series $\{x_t; t = 0, \pm 1, \cdots\}$ is **ARMA**(p, q) if it is stationary and

$$x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \cdots + \theta_q w_{t-q} \quad (27)$$

  where $w_t \sim wn(0, \sigma_w^2)$, and $\phi_p \neq 0, \theta_q \neq 0$, and $\sigma_w^2 > 0$.
- The parameters p and q are called the autoregressive and the moving average orders, respectively.
- If x_t has a nonzero mean μ, we set $\alpha = \mu(1 - \phi_1 - \cdots - \phi_p)$ and have

$$x_t = \alpha + \phi_1 x_{t-1} + \phi_2 x_{t-2} + \cdots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \cdots + \theta_q w_{t-q} \quad (28)$$

- With autoregressive and moving average operators, **ARMA**(p, q) model is:

$$\phi(B) x_t = \theta(B) w_t \quad (29)$$

Autoregressive Moving Average Models

- There are following problems for **ARMA**(p, q)
  1. parameter redundant models,
  2. stationary AR models that depend on the future, and
  3. MA models that are not unique.
- To overcome these problems, we will require some additional restrictions on the model parameters.
- The **AR and MA polynomials** are defined as

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p, \quad \phi_p \neq 0 \tag{30}$$

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q, \quad \theta_q \neq 0 \tag{31}$$

  respectively, where z is a complex number.
- We require that $\phi(z)$ and $\theta(z)$ have no common factors.

Autoregressive Moving Average Models

- An $ARMA(p,q)$ model is said to be **causal**, if the time series $\{x_t; t = 0, \pm 1, \cdots\}$ can be written as a one-sided linear process

$$x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j} = \psi(B) w_t \tag{32}$$

where $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$, and $\sum_{j=0}^{\infty} |\psi_j| < \infty$; we set $\psi_0 = 1$.

- An $ARMA(p,q)$ model is causal if and only if $\phi(z) \neq 0$ for $|z| \leq 1$. The coefficients of the linear process can be determined by solving

$$\psi(z) = \sum_{j=0}^{\infty} \psi_j z^j = \frac{\theta(z)}{\phi(z)}, \quad |z| \leq 1 \tag{33}$$

Autoregressive Moving Average Models

- An $ARMA(p, q)$ model is said to be **invertible**, if the time series $\{x_t; t = 0, \pm 1, \cdots\}$ can be written as

$$\pi(B)x_t = \sum_{j=0}^{\infty} \pi_j x_{t-j} = w_t \tag{34}$$

where $\pi(B) = \sum_{j=0}^{\infty} \pi_j B^j$, and $\sum_{j=0}^{\infty} |\pi_j| < \infty$; we set $\pi_0 = 1$.

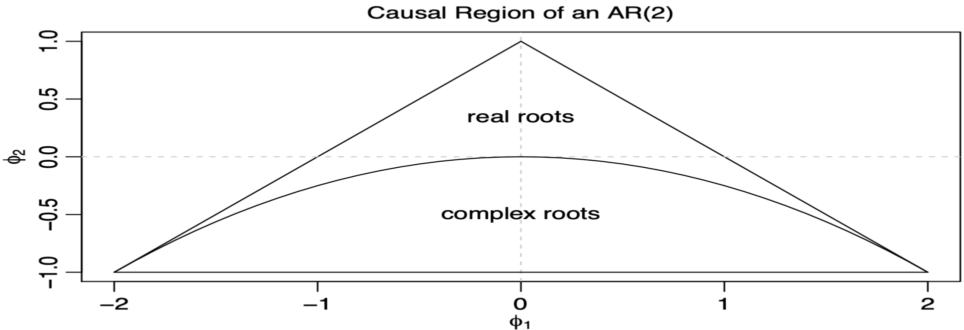
- An $ARMA(p, q)$ model is invertible if and only if $\theta(z) \neq 0$ for $|z| \leq 1$. The coefficients of $\pi(B)$ can be determined by solving

$$\pi(z) = \sum_{j=0}^{\infty} \pi_j z^j = \frac{\phi(z)}{\theta(z)}, \quad |z| \leq 1 \tag{35}$$

Autoregressive Moving Average Models

- For an AR(1) model, $(1 - \phi B)x_t = w_t$, to be causal, the root of $\phi(z) = 1 - \phi z$ must lie outside of the unit circle. That is, $|\phi| < 1$.
- Consider the AR(2) model, $(1 - \phi_1 B - \phi_2 B^2)x_t = w_t$. the causal condition requires that the two roots of $\phi(z) = 1 - \phi_1 z - \phi_2 z^2$ lie outside of the unit circle. That is $\left|\frac{\phi_1 \pm \sqrt{\phi_1^2 + 4\phi_2}}{-2\phi_2}\right| > 1$, which is equivalent to

$$\phi_1 + \phi_2 < 1, \quad \phi_2 - \phi_1 < 1, \quad |\phi_2| < 1 \tag{36}$$



Difference Equations

- Recall that ACF of AR(1), $\rho(h)$, satisfies the recursion $\rho(h) = \phi\rho(h-1)$.
- In general, the *homogeneous difference equation of order 1*

$$u_n - \alpha u_{n-1} = 0, \quad \alpha \neq 0, \; n = 1, 2, \cdots \tag{37}$$

has the solution $u_n = \alpha^n c$ for initial condition $u_0 = c$.

- This can also be written as

$$u_n = \alpha^n c = (z_0^{-1})^n c \tag{38}$$

with $z_0 = 1/\alpha$ being the root of the characteristic polynomial $\alpha(z) = 1 - \alpha z$.

Difference Equations

- The *homogeneous difference equation of order 2*

$$u_n - \alpha_1 u_{n-1} - \alpha_2 u_{n-2} = 0, \quad \alpha_2 \neq 0, \; n = 2, 3, \cdots \tag{39}$$

has characteristic polynomial $\alpha(z) = 1 - \alpha_1 z - \alpha_2 z^2$ with two roots z_1, z_2.

- If $z_1 \neq z_2$, the solution of the difference equation has the following format

$$u_n = c_1 z_1^{-n} + c_2 z_2^{-n} \tag{40}$$

where c_1 and c_2 can be determined by two initial conditions u_0 and u_1.

- If $z_1 = z_2 := z_0$, then the solution is

$$u_n = z_0^{-n}(c_1 + c_2 n) \tag{41}$$

where c_1 and c_2 can also be determined by two initial conditions u_0 and u_1.

Difference Equations

- In general, the *homogeneous difference equation of order p*

$$u_n - \alpha_1 u_{n-1} - \cdots - \alpha_p u_{n-p} = 0, \quad \alpha_p \neq 0, \; n = p, p+1, \cdots \tag{42}$$

  has characteristic polynomial $\alpha(z) = 1 - \alpha_1 z - \cdots - \alpha_p z^p$.
- Suppose $\alpha(z)$ has r distinct roots, $z_1, \cdots, z_r$ with multiplicities $m_1, \cdots, m_r$ respectively ($\sum_{j=1}^r m_j = p$). Then general solution is

$$u_n = z_1^{-n} P_1(n) + \cdots + z_r^{-n} P_r(n) \tag{43}$$

  where $P_j(n)$, for $j = 1, \cdots, r$, is a polynomial of n, of degree $m_j - 1$, and can be solved jointly by initial conditions $u_0, \cdots, u_{p-1}$.
- How does it apply to obtain the ACF for AR(p), e.g. AR(2)?

Difference Equations

- Recall that we could use matching coefficients to solve ARMA(p,q) model $\phi(B)x_t = \theta(B)w_t$ and write $x_t = \psi(B)w_t = \sum_{j=0}^{\infty}\psi_j w_{t-j}$.
- Then by matching coefficients in $\phi(z)\psi(z) = \theta(z)$ we get

$$\begin{aligned} \psi_0 &= 1 \\ \psi_1 - \phi_1\psi_0 &= \theta_1 \\ \psi_2 - \phi_1\psi_1 - \phi_2\psi_0 &= \theta_2 \quad \cdots \end{aligned}$$

  where we should take $\phi_j = 0$ for $j > p$ and $\theta_j = 0$ for $j > q$.
- Then the ψ-weights satisfy the homogeneous difference equation

$$\psi_j - \sum_{k=1}^{p}\phi_k\psi_{j-k} = 0, \quad j \geq \max(p, q+1) \tag{44}$$

  with initial conditions

$$\psi_j - \sum_{k=1}^{p}\phi_k\psi_{j-k} = \theta_j, \quad 0 \leq j < \max(p, q+1) \tag{45}$$

Autocorrelation

- First, recall the autocovariance of an MA(q) process. $x_t = \theta(B) w_t$ can be obtained

$$\gamma(h) = \mathrm{Cov}(x_{t+h}, x_t) = \begin{cases} \sigma_w^2 \sum_{j=0}^{q-h} \theta_j \theta_{j+h}, & 0 \le h \le q \\ 0, & h > q \end{cases} \tag{46}$$

which implies the ACF

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \begin{cases} \frac{\sum_{j=0}^{q-h} \theta_j \theta_{j+h}}{\sum_{j=0}^{q} \theta_j^2}, & 1 \le h \le q \\ 0, & h > q \end{cases} \tag{47}$$

- Then, consider the general ARMA(p,q) model, the autocovariance function can be obtained $\gamma(h) = \mathrm{Cov}(x_{t+h}, x_t) = \sigma_w^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h}$ with ψ-weights.

Autocorrelation

- Alternatively, we could use the following difference equation

$$\gamma(h) = \mathrm{Cov}\left(\sum_{j=1}^{p}\phi_j x_{t+h-j} + \sum_{j=0}^{q}\theta_j w_{t+h-j}, x_t\right) = \sum_{j=1}^{p}\phi_j\gamma(h-j) + \sigma_w^2\sum_{j=h}^{q}\theta_j\psi_{j-h} \tag{48}$$

- This yields a *general homogeneous equation for the ACF of a causal ARMA process*

$$\gamma(h) - \sum_{j=1}^{p}\phi_j\gamma(h-j) = 0, \quad h \geq \max(p, q+1) \tag{49}$$

with initial conditions

$$\gamma(h) - \sum_{j=1}^{p}\phi_j\gamma(h-j) = \sigma_w^2\sum_{j=h}^{q}\theta_j\psi_{j-h}, \quad 0 \leq h \leq \max(p, q+1) \tag{50}$$

Autocorrelation

Example

How to obtain the ACF of ARMA(1,1) process, $x_t = \phi x_{t-1} + w_t + \theta w_{t-1}$, *with* $|\phi| < 1$*?*

Partial Autocorrelation

- ACF provides a considerable amount of information about the order of the dependence for MA(q). However, the ACF alone tells us little about the orders of dependence for AR(p) or ARMA(p,q).
- The **partial autocorrelation function (PACF)** of a stationary process, x_t, denoted ϕ_{hh}, for $h = 1, 2, \cdots$ is

$$\phi_{hh} = \begin{cases} \operatorname{corr}(x_{t+1}, x_t) = \rho(1), & h = 1 \\ \operatorname{corr}(x_{t+h} - \hat{x}_{t+h}, x_t - \hat{x}_t), & h \geq 2 \end{cases} \tag{51}$$

  where we have $\hat{x}_{t+h} = \sum_{j=1}^{h-1} \beta_j x_{t+h-j}$ and $\hat{x}_t = \sum_{j=1}^{h-1} \beta_j x_{t+j}$.
- PACF, ϕ_{hh}, is the correlation between x_{t+h} and x_t with the linear dependence of $\{x_{t+1}, \cdots, x_{t+h-1}\}$ on each, removed.
- If x_t is Gaussian, then $\phi_{hh} = \operatorname{corr}(x_{t+h}, x_t | x_{t+1}, \cdots, x_{t+h-1})$.
- PACF cuts off after lag p for AR(p), i.e. $\phi_{hh} = 0$ for $h > p$.

Forecasting

- In forecasting, the goal is to predict future values of a time series, x_{n+m}, $m = 1, 2, \cdots$, based on the data collected to the present, $x_{1:n} = \{x_1, \cdots, x_n\}$.
- It can be shown that the minimum mean square error predictor of x_{n+m} is

$$x_{n+m}^n = \mathrm{E}[x_{n+m}|x_{1:n}] \tag{52}$$

- We restrict to predictors that are linear functions of the data, that is,

$$x_{n+m}^n = \alpha_0 + \sum_{k=1}^n \alpha_k x_k \tag{53}$$

- The **Best Linear Predictor (BLP)** for stationary process x_t is found by solving

$$\mathrm{E}[(x_{n+m} - x_{n+m}^n)x_k] = 0, \quad k = 0, 1, \cdots, n \tag{54}$$

where $x_0 = 1$ for $\alpha_0, \alpha_1, \cdots, \alpha_n$.

Forecasting

- First, consider *one-step-ahead prediction*. $x_{n+1}^n = \sum_{j=1}^n \phi_{nj} x_{n+1-j}$. The BLP satisfies

$$\sum_{j=1}^{n} \phi_{nj}\gamma(k-j) = \gamma(k), \quad k = 1, \cdots, n \tag{55}$$

- This prediction can be written in matrix notation

$$\Gamma_n \phi_n = \gamma_n \tag{56}$$

where $\Gamma_n = \{\gamma(k-j)\}_{j,k=1}^n$, $\phi_n = (\phi_{n1}, \cdots, \phi_{nn})'$, $\gamma_n = (\gamma(1), \cdots, \gamma(n))'$.

- For ARMA models, we have $\sigma_w^2 > 0$, and $\gamma(h) \to 0$ as $h \to \infty$. Thus Γ_n is positive definite. The one-step-ahead BLP, x_{n+1}^n, is solved as

$$x_{n+1}^n = \phi_n' x, \quad \phi_n = \Gamma_n^{-1}\gamma_n \tag{57}$$

- The *mean square one-step-ahead prediction error* is

$$P_{n+1}^n = \mathrm{E}(x_{n+1} - x_{n+1}^n)^2 = \gamma(0) - \gamma_n'\Gamma_n^{-1}\gamma_n \tag{58}$$

Forecasting

Example

Consider one-step-ahead prediction of AR(2) model, $x_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + w_t$.

Estimation

- Given a process, x_t, how do we determine p, q and $\phi_1, \cdots, \phi_p$, $\theta_1, \cdots, \theta_q$ if we want to model it using ARMA(p,q)?
- We can use *method of moments* estimators. Consider AR(p) first:

$$x_t = \sum_{j=1}^{p} \phi_j x_{t-j} + w_t \tag{59}$$

- Recall the first $p+1$ (difference) equations (49)(50) for ACF of ARMA, which defines the following **Yule-Walker equations**

$$\gamma(h) = \sum_{j=1}^{p} \phi_j \gamma(h-j) = 0, \quad h = 1, 2, \cdots, p \tag{60}$$

$$\sigma_w^2 = \gamma(0) - \sum_{j=1}^{p} \phi_j \gamma(j) = 0. \tag{61}$$

Estimation

- In matrix notation, the Yule-Walker equations are

$$\Gamma_p \phi = \gamma_p, \quad \sigma_w^2 = \gamma(0) - \phi' \gamma_p \tag{62}$$

where $\Gamma_p = \{\gamma(k-j)\}_{j,k=1}^p$, $\phi = (\phi_1, \cdots, \phi_p)'$, $\gamma_p = (\gamma(1), \cdots, \gamma(p))'$.

- Using the method of moments, we replace $\gamma(h)$ by $\hat{\gamma}(h)$ and solve

$$\hat{\phi} = \hat{\Gamma}_p^{-1} \hat{\gamma}_p, \quad \hat{\sigma}_w^2 = \hat{\gamma}(0) - \hat{\gamma}_p' \hat{\Gamma}_p^{-1} \hat{\gamma}_p \tag{63}$$

- Sometimes it is more convenient to work with the sample ACF so the Yule-Walker estimator can be written as

$$\hat{\phi} = \hat{R}_p^{-1} \hat{\rho}_p, \quad \hat{\sigma}_w^2 = \hat{\gamma}(0)[1 - \hat{\rho}_p' \hat{R}_p^{-1} \hat{\rho}_p] \tag{64}$$

where $\hat{R}_p = \{\hat{\rho}(k-j)\}_{j,k=1}^p$, and $\hat{\rho}_p = (\hat{\rho}(1), \cdots, \hat{\rho}(p))'$.

Table of Contents

Integrated Models for Nonstationary Data

- Recall that non-stationary time series data can be modeled as a composition of nonstationary trend and a zero-mean stationary component

$$x_t = \mu_t + y_t \tag{65}$$

- In many cases (linear drift model), differencing can remove the trend and render a stationary residual process

$$\nabla x_t = v_t + \nabla y_t \tag{66}$$

where $\nabla = 1 - B$, and v_t is stationary, e.g. $\mu_t = \mu_{t-1} + v_t$.

- When μ_t is a k-th order polynomial, $\mu_t = \sum_{j=1}^{k} \beta_j t^j$, $\nabla^k x_t$ is stationary.

Integrated ARMA (ARIMA) models

- A process x_t is said to be **ARIMA(p, d, q)** if

$$\nabla^d x_t = (1 - B)^d x_t \tag{67}$$

  is ARMA(p, q).
- In general, we will write the model as

$$\phi(B)(1 - B)^d x_t = \theta(B) w_t \tag{68}$$

- If $\mathrm{E}(\nabla^d x_t) = \mu$, we write the model as

$$\phi(B)(1 - B)^d x_t = \delta + \theta(B) w_t \tag{69}$$

  where $\delta = \mu(1 - \sum_{j=1}^p \phi_j)$.

Integrated ARMA (ARIMA) models

- Since $y_t = \nabla^d x_t$ is ARMA(p,q), previous theories/methods for ARMA models apply.
- For example, if $d = 1$, given forecasts y^n_{n+m} for $m = 1, 2, \cdots$, we have $y^n_{n+m} = \nabla^d x^n_{n+m}$ such that

$$x^n_{n+m} = y^n_{n+m} + x^n_{n+m-1} \tag{70}$$

  with initial condition $x^n_{n+1} = y^n_{n+1} + x_n$ (noting $x^n_n = x_n$).
- The mean-squared prediction error, P^n_{n+m}, can be approximated by

$$P^n_{n+m} = \sigma^2_w \sum_{j=0}^{m-1} \psi^{*2}_j \tag{71}$$

  where ψ^*_j is the coefficient of z_j in $\psi^*(z) = \theta(z)/\phi(z)(1-z)^d$.

Integrated ARMA (ARIMA) models

- Consider the random walk with drift model, $x_t = \delta + x_{t-1} + w_t$ $(x_0 = 0)$, which can be recognized as a trivial ARIMA(0,1,0).
- Given data $x_{1:n}$, the one-step- ahead forecast is given by

$$x^n_{n+1} = \mathrm{E}[x_{n+1}|x_{1:n}] = \mathrm{E}[\delta + x_n + w_{n+1}|x_{1:n}] = \delta + x_n \tag{72}$$

- The two-step-ahead forecast is given by $x^n_{n+2} = \delta + x^n_{n+1} = 2\delta + x_n$, and consequently, the m-step-ahead forecast is

$$x^n_{n+m} = m\delta + x_n \tag{73}$$

- Note we can write $x_{n+m} = (n+m)\delta + \sum_{j=1}^{n+m} w_j = m\delta + x_n + \sum_{j=n+1}^{n+m} w_j$.
- The m-step-ahead prediction error is given

$$P^n_{n+m} = \mathrm{E}(x_{n+m} - x^n_{n+m})^2 = \mathrm{E}\left(\sum_{j=n+1}^{n+m} w_j\right)^2 = m\sigma_w^2 \tag{74}$$

Integrated ARMA (ARIMA) models

Example

Consider ARIMA(0,1,1), IMA(1,1) model:

$$x_t = x_{t-1} + w_t - \lambda w_{t-1} \tag{75}$$

Show that

$$x_t = \sum_{j=1}^{\infty} (1-\lambda)\lambda^{j-1} x_{t-j} + w_t \tag{76}$$

Integrated ARMA (ARIMA) models

There are a few basic steps to fitting ARIMA models to time series data.

1. plotting the data,
2. possibly transforming the data,
3. identifying the dependence orders of the model,
4. parameter estimation,
5. diagnostics, and
6. model choice.