Lecture 14: Autoregressive integrated Moving Average model (1)

Wengang Mao (Marine Technology)
Department of Mechanics and Maritime Sciences,
Chalmers University of Technology,
Goteborg, Sweden

Outline of ARIMA models

- **Basis of autoregressive integrated moving average (ARIMA) models**
- **Difference equations**
- **Autocorrelation (ACF) and partial autocorrelation (PACF)**
- **Forecasting**
- Estimation
- Integrated Models for Nonstationary Data
- Steps to build ARIMA Models

Basis of ARMA models

- **Let denote a time series X_t, a series of dependent random variables as:**
  - $X_t = [X_1, X_2, X_3, \dots, X_{t-1}, X_t]$
  - The random variables are correlated: autocorrelation function (ACF)
  - Be 2nd order stationary: constant mean and ACF only dependent on time lag

  - From a non-stationary process to a stationary process
    - ✓ *Detrend (subtract the mean trend from the original process/time series)*
    - ✓ *Difference (of different orders)*
    - ✓ *Transform the time series/process, such as logarithm, or exponential transform*
- **ACF for AR (autoregressive) and ARMA (autoregressive moving average) models**
- **Adding non-stationarity leads to Auto-Regressive Integrated Moving Average (ARIMA)**

ARIMA: autoregressive AR models

- AR models: current value of X_t can be expressed by previous values
- An ***autoregressive model*** *of order p*, **AR**(*p*) can be expressed by:
  - ✓$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} +, \dots, +\phi_p X_{t-p} + W_t$
  - ✓X_t is stationary (otherwise detrend/difference/transform the original process),
  - ✓$W_t \sim wn(0, \sigma_w^2)$, i.e., a white noise process (stationary)
  - ✓$\phi_1, \phi_2, \dots, \phi_p$ are constants while $\phi_p \neq 0$.
- The above AR model can be rewrite by the backshift operator as
  - ✓$\left(1 - \phi_1 B - \phi_2 B^2 -, \dots, -\phi_p B^p\right) X_t = W_t$
  - ✓Or more concisely,$\phi(B) = W_t$
  - ✓The properties of $\phi(B)$ are important to define the AR model
- Generally, the ***Autoregressive Operator*** is defined by
  - ✓$\phi(B) = 1 - \phi_1 B - \phi_2 B^2 -, \dots, -\phi_p B^p$

ARIMA: AR(1) models

- AR model is often to model the stationary time series/random process (If there is a trend, it should first be removed by, detrending, difference, or transform, to form a stationary process)
- The AR(1) model is denoted by: $X_t = \phi X_{t-1} + W_t$ (or alternatively $(1-\phi B)X_t = W_t$)
- If $|\phi| < 1$ and $var(X_t) < \infty$, the AR(1) models can be represented **by a linear process** as
  - ✓$X_t = \phi X_{t-1} + W_t = \phi(\phi X_{t-2} + W_{t-1}) + W_t =, \dots, = \phi^k X_{t-k} + \sum_{j=0}^{k-1} \phi^j W_{t-j}$, since $|\phi| < 1$, then
  - ✓$X_t = \sum_{j=0}^{\infty} \phi^j W_{t-j}$.
  - ✓To prove the stationarity of AR(1) with $|\phi| < 1$, we show the first two moments of the process
  - ✓$E[X_t] = \sum_{j=0}^{\infty} \phi^j E[W_{t-j}] = 0$
  - ✓The covariance function should be only dependent on the time lag as

$$\begin{aligned}\gamma(h) = \operatorname{cov}(x_{t+h}, x_t) &= \mathrm{E}\left[\left(\sum_{j=0}^{\infty} \phi^j w_{t+h-j}\right)\left(\sum_{k=0}^{\infty} \phi^k w_{t-k}\right)\right] \\ &= \mathrm{E}\left[\left(w_{t+h} + \cdots + \phi^h w_t + \phi^{h+1} w_{t-1} + \cdots\right)(w_t + \phi w_{t-1} + \cdots)\right] \\ &= \sigma_w^2 \sum_{j=0}^{\infty} \phi^{h+j} \phi^j = \sigma_w^2 \phi^h \sum_{j=0}^{\infty} \phi^{2j} = \frac{\sigma_w^2 \phi^h}{1-\phi^2}, \quad h \geq 0.\end{aligned}$$

ARIMA: AR(1) examples (1)

Let take a look at how the AR(1) process looks like

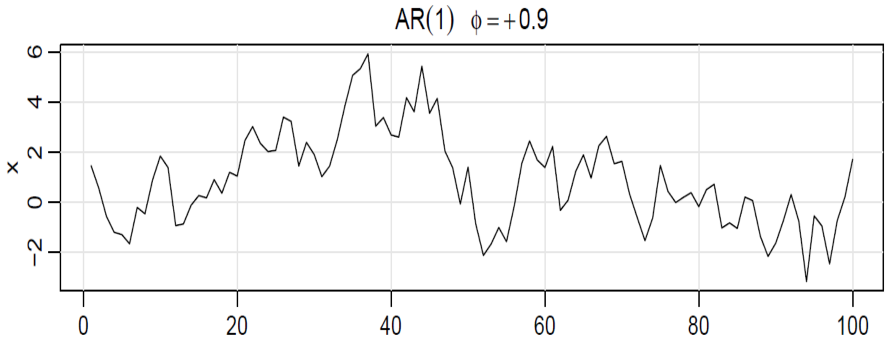
✓$X_t = \phi X_{t-1} + W_t$, where $W_t \sim wn(0,1)$

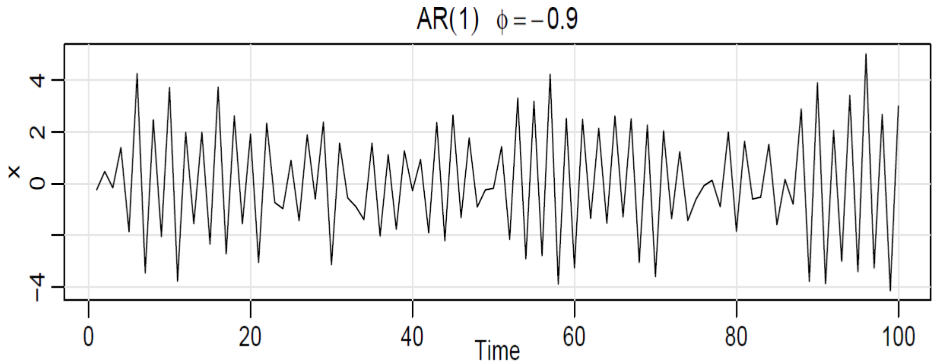
✓The ACF is computed as:

$$\rho(h) = \rho(X_t, X_{t-h}) = \rho\left(\phi^h X_{t-h} + \sum_{j=0}^{h-1} \phi^j W_{t-j}, X_{t-h}\right) = \rho\left(\phi^h X_{t-h}, X_{t-h}\right) + \rho\left(\sum_{j=0}^{h-1} \phi^j W_{t-j}, X_{t-h}\right) = \phi^h$$

✓If $\phi > 0$, continuous time points are positively correlated, i.e., adjacent points close to each other. It means the time series will be relatively smooth

✓If $\phi < 0$, adjacent points behave interactively change between positive and negative correlated.





ARIMA: AR(1) examples (2)

- Properties of the AR(1) process: $X_t = \phi X_{t-1} + W_t$
  - ✓$|\phi| < 1$ is often a **causal process**, i.e., one may know the current value through previous values
  - ✓$|\phi| \geq 1$ is often an **explosive process**, because the values quickly become large in magnitude
- An explosive process can be converted into a causal process by:
  - ✓$X_t = \phi^{-1} X_{t+1} - \phi^{-1} W_{t+1}$, by iterating forward many steps, the AR(1) becomes
  - ✓$X_t = -\sum_{j=0}^{\infty} \phi^{-j} W_{t+j}$ a linear process (because $|\phi|^{-1} < 1$)
  - ✓The above case is stationary, also future dependent, but not **causal.**

ARIMA: AR(1) examples (3)

- $X_t = \phi X_{t-1} + W_t$, with $|\phi| > 1$ and $W_t \sim iid\ N(0, \sigma_w^2)$
  - ✓X_t is a non-causal stationary Gaussian process with $E[X_t] = 0$

$$\gamma_x(h) = \operatorname{cov}(x_{t+h}, x_t) = \operatorname{cov}\left(-\sum_{j=1}^{\infty} \phi^{-j} w_{t+h+j}, -\sum_{k=1}^{\infty} \phi^{-k} w_{t+k}\right)$$
$$= \sigma_w^2 \phi^{-2} \phi^{-h} / (1 - \phi^{-2}).$$

  - ✓For example, $X_t = 2X_{t-1} + W_t$ with $\sigma_w^2 = 1 \Leftrightarrow Y = 0.5Y_{t-1} + V_t$ with $\sigma_V^2 = 1/4$
  - ✓The non-causal stationary property contains only future observations, from statistical point, it does not make sense → non-stationary process.

ARIMA: AR(1) examples (4)

For the AR(1) $\phi(B)X_t = W_t$, it can be written as a linear process

- $X_t = \sum_{j=0}^{\infty} \phi^j W_{t-j} = \sum_{j=0}^{\infty} \psi_j W_{t-j} = \psi(B) W_t$:
- We have $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$

Let inject $X_t = \psi(B)W_t$ into $\phi(B)X_t = W_t \Rightarrow \phi(B)\psi(B)W_t = W_t$

- $\phi(B)\psi(B) = 1 \Rightarrow (1 - \phi B)\left(1 + \psi_1 B + \psi_2 B^2 + \cdots + \psi_j B^j + \cdots\right) = 1 \Rightarrow$
- $1 + (\psi_1 - \phi)B + (\psi_2 - \psi_1\phi)B^2 + \cdots + \left(\psi_j - \psi_{j-1}\phi\right)B^j + \cdots = 1 \Rightarrow$
- $\psi_j = \psi_{j-1}\phi$, with $\psi_0 = 1$ we can derive $\psi_j = \phi^j$

This will help represent AR(p) by the moving average MA(q) model.

ARIMA: MA(q) models

- ***The moving average model of order q, MA(q), is defined by***
  - ✓$X_t = W_t + \theta_1 W_{t-1} + \theta_2 W_{t-2} + \cdots + \theta_q W_{t-q}$
  - ✓$W_t \sim wn(0, \sigma_W^2)$ and θ_1, θ_2, ..., θ_q ($\theta_q \neq 0$) are the parameters to define MA(q)
  - ✓The MA formula is similar as the linear process of AR(1) with $\psi_0 = 1$ and $\theta_j = \psi_j$
  - ✓Let use backshift to write the MA: $X_t = W_t + \theta_1 B W_t \; + \theta_2 B^2 W_t + \cdots + \theta_q B^q W_t$
  - ✓Then **MA**(q) can be written as $X_t = \theta(B) W_t$
- ***The moving average operator is then defined as***
  - ✓$\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + +\theta_q B^q$
  - ✓The MA model is stationary for any values of the parameters θ_1, θ_2, ..., θ_q
  - ✓The key to define a MA process is to 1) get the order q, and 2) get the parameters

ARIMA: MA(1) model examples

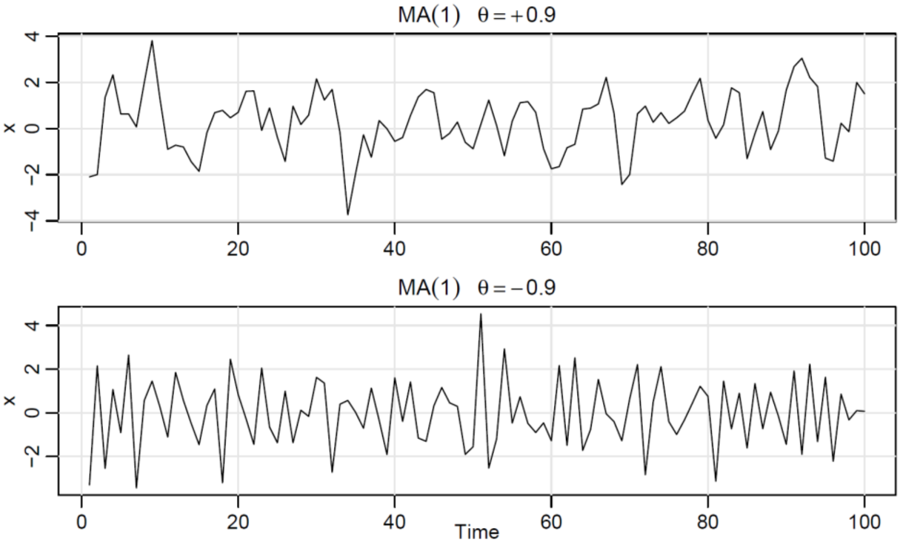
- ***The MA(1) process as defined by $X_t = W_t + \theta W_{t-1}$.***
  - ✓$E[X_t] = 0$
  - ✓The autocovariance function can be written as

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

  - ✓The autocorrection function (ACF) then becomes

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

  - ✓In MA(1) process, X_t is only correlated with X_{t-1}, not X_{t-2}, others
  - ✓In the AR(1) process, the correlation $\rho(X_t, X_{t-k}) \neq 0$




```
par(mfrow = c(2,1))
plot(arima.sim(list(order=c(0,0,1), ma=.9), n=100), ylab="x",
        main=(expression(MA(1)~~~theta==+.5)))
plot(arima.sim(list(order=c(0,0,1), ma=-.9), n=100), ylab="x",
        main=(expression(MA(1)~~~theta==-.5)))
```

ARIMA: MA(1) model non-uniqueness and invertibility

- **The following MA(1) processes have the same model properties**

$$x_t = w_t + \tfrac{1}{5} w_{t-1}, \quad w_t \sim \text{iid } \mathrm{N}(0, 25)$$

$$y_t = v_t + 5 v_{t-1}, \quad v_t \sim \text{iid } \mathrm{N}(0, 1)$$

$$\gamma(h) = \begin{cases} 26 & h = 0, \\ 5 & h = 1, \\ 0 & h > 1. \end{cases}$$

- For convenience, by mimicking the criterion of causality for AR models, we will choose the model with an infinite AR representation. Such a process is called an ***invertible*** process.

ARIMA: ARMA models

- A time series $X_t: t = 0, \pm 1, \pm 2, \ldots$ is **ARMA**(*p*, *q*) if it is stationary and
  - ✓ $X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + W_t + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q}$
  - ✓ $\phi_p \neq 0$, $\theta_q \neq 0$ and $\sigma_W^2 \geq 0$
  - ✓ The parameters *p* and *q* are called the AR and the MA orders.
  - ✓ If $E[X_t] \neq 0$, take away the mean/trend to form the stationary ARMA process
  - ✓ The ARMA can be written in backshift method as: $\phi(B)X_t = \theta(B)W_t$
  - ✓ **A potential problem:** the model can be unnecessarily complex by adding any arbitrary operator $\eta(B)$, $\eta(B)\phi(B)X_t = \eta(B)\theta(B)W_t$

- Model/parameter redundancy (an example)
  - ✓ Consider a white noise model as: $X_t = W_t$. If it is multiplied by $\eta(B) = 1 - 0.5B$ on both sides, then
  - ✓ $(1 - 0.5B)X_t = (1 - 0.5B)W_t \Rightarrow X_t = 0.5X_{t-1} - 0.5W_{t-1} + W_t$, ARMR(1,1) model, X_t is still a white noise

ARMA: problems and further restrictions

- In the process to construct a **ARMA**(p,q) model, there might be three potential problems

   (i) parameter redundant models,
   (ii) stationary AR models that depend on the future, and
   (iii) MA models that are not unique.

- Additional restrictions should be introduced to address the above problems.
- First, we need to introduce further definitions

The **AR and MA polynomials** *are defined as*

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

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

ARMA: problems and further restrictions

- To problem 1: parameter redundant, i.e., **ARMA**(p,q) in its simplest form, it requires *the AR and MA polynomials* $\phi(z)$ *and* $\theta(z)$ *have no common factors*.

- To problem 2: future dependent model, we can introduce the concept of *causality*, viz,
  - ✓An **ARMA**(*p, q*) model is causal if X_t can be written as a one-side linear process:
  - ✓$X_t = \sum_{j=0}^{\infty} \psi_j W_{t-j} = \psi(B) W_t$, where $\psi(B) = \sum_{j=0}^{\infty} \psi_j B^j$, $\sum_{j=0}^{\infty} |\psi_j| < \infty$ while $\psi_0 = 1$
  - ✓Alternatively, **ARMA**(*p, q*) model is *causal* if and only if $\phi(z) \neq 0$ for $|z| \leq 1$. Then, the coefficients in the above linear process can be estimated by:

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

  - ✓Another alternative is that **ARMA**(*p, q*) model is *causal* if and only if when the roots of $\phi(z)$ lie outside the unit circle, i.e., $\phi(z) = 0$ only when $|z| > 1$.

ARMA: problems and further restrictions

- To the problem 3: uniqueness of MA, the model with infinite autoregressive representation should be selected.
- **Invertibility** is often used to define the uniqueness of a ARMA model.
- There are two possibilities to check the **invertibility**.
  - ✓ An **ARMA**(*p*, *q*) model is invertible, if the time series X_t can be written as
  - ✓ $\pi(B)X_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} = W_t$, where $\pi(B) = \sum_{j=0}^{\infty} \pi_j B^j$ and $\sum_{j=0}^{\infty} |\pi_j| < \infty$.
  - ✓ Similar as the parameter redundant, an **ARMA**(*p*, *q*) model is *invertible* if and only if $\theta(z) \neq 0$ for $|z| \leq 1$. Then, the coefficients in the above process can be estimated by:

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

  - ✓ Another alternative is that **ARMA**(*p*, *q*) model is *invertible* if and only if when the roots of $\theta(z)$ lie outside the unit circle, i.e., $\theta(z) = 0$ only when $|z| > 1$.

ARMA: restrictions and examples

Parameter Redundancy, Causality, Invertibility of ARMA(*p*, *q*) models

- ✓Let $X_t = 0.4X_{t-1} + 0.45X_{t-2} + W_t + W_{t-1} + 0.25W_{t-2}$
- ✓In a backshift form, $(1 - 0.4B - 0.45B^2)X_t = (1 + B + 0.25B^2)W_t$, an **ARMA**(2,2) model
- ✓Notice $\phi(B) = (1 + 0.5B)(1 - 0.9B)$ & $\theta(B) = (1 + 0.5B)^2$, with a common factor ➔
- ✓A simplified new ARMA model:

  $X_t = 0.9X_{t-1} + 0.5W_{t-1} + W_t$ (no parameter redundancy)
- ✓The model is *causal* because $\phi(z) = (1 - 0.9z) = 0$ when $z = \frac{10}{9}$ outside the circle.
- ✓The model is *invertible* because $\theta(z) = (1 + 0.5z) = 0$ when $z = -2$ outside the circle

ARMA: restrictions and examples

To write the above **ARMA** ($X_t = 0.9X_{t-1} + 0.5W_{t-1} + W_t$) into a linear process.

- For the AR model, i.e., $X_t = \sum_{j=0}^{\infty} \psi_j W_{t-j} = \psi(B)W_t$ (from **ARMA** to **MA**)
  - ✓We need to set up the relationship $\phi(z)\psi(z) = \theta(z)$ ➔ $(1 - .9z)(1 + \psi_1 z + \psi_2 z^2 + \cdots + \psi_j z^j + \cdots) = 1 + .5z.$

$$1 + (\psi_1 - .9)z + (\psi_2 - .9\psi_1)z^2 + \cdots + (\psi_j - .9\psi_{j-1})z^j + \cdots = 1 + .5z.$$

$$x_t = w_t + 1.4 \sum_{j=1}^{\infty} .9^{j-1} w_{t-j}$$

```
ARMAtoMA(ar = .9, ma = .5, 10)   # first 10 psi-weights
[1] 1.40 1.26 1.13 1.02 0.92 0.83 0.74 0.67 0.60 0.54
```

- For the MA model, i.e., $\pi(B)X_t = \sum_{j=0}^{\infty} \pi_j X_{t-j} = W_t$
  - ✓We need to set up the relationship $\theta(z)\psi(z) = \phi(z)$

$$(1 + .5z)(1 + \pi_1 z + \pi_2 z^2 + \pi_3 z^3 + \cdots) = 1 - .9z$$

$$x_t = 1.4 \sum_{j=1}^{\infty} (-.5)^{j-1} x_{t-j} + w_t$$

```
ARMAtoMA(ar = -.5, ma = -.9, 10)   # first 10 pi-weights
[1] -1.400 .700 -.350 .175 -.087 .044 -.022 .011 -.006 .003
```

Outline of ARIMA models

- Basis of autoregressive moving average models
- Difference equations
- Autocorrelation (ACF) and partial autocorrelation (PACF)
- Forecasting
- Estimation
- Integrated Models for Nonstationary Data
- Multiplicative Seasonal ARIMA Models

Difference equation

Difference equation: describe/estimate the ACF of a ARMA process

- An *n*th order linear difference equation can be written in terms of $a_1, \ldots, a_n$ and b as

$$y_t = a_1 y_{t-1} + \cdots + a_n y_{t-n} + b,$$

- The equation is called *homogeneous* if $b = 0$ and *nonhomogeneous* if $b \neq 0$.

- The ***ACF*** of an AR(2) process $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + W_t$

  ✓The AR(2) associated polynomial: $\phi(z) = 1 - \phi_1 z - \phi_2 z^2$

  ✓This AR(2) is an ARMA □ AR(2) is causal ➔ the roots of above polynomial $|z| > 1$ (causal stationary)

  ✓$E[X_t X_{t-h}] = \phi_1 E[X_{t-1} X_{t-h}] + \phi_2 E[X_{t-2} X_{t-h}] + E[W_t X_{t-h}]$

  ✓$\gamma(h) = \phi_1 \gamma(h-1) + \phi_2 \gamma(h-2)$

  ✓The ACF: $\rho(h) = \phi_1 \rho(h-1) + \phi_2 \rho(h-2)$, since $\rho(0) = 1$, ➔ $\rho(1) = \phi_1 / (1 - \phi_2)$

  ✓$\rho(2) = \phi_1 \rho(2-1) + \phi_2 \rho(2-2) \rightarrow \rho(2) = \phi_1 \rho(1) + \phi_2, \ldots$

Difference equation: ψ-weight ARMA

Let an **ARMA**(*p*, *q*) process denoted by $\phi(B)X_t = \theta(B)W_t$

- ✓According to causality of an ARMA, we have $X_t = \sum_{j=0}^{\infty} \psi_j W_{t-j}$
- ✓For a pure **MA**(*q*) model, ψ-weights are easily obtained as $\psi_0 = 1$ and $\psi_j = \theta_j$ for $j \leq q$, and 0 others
- ✓For a general **ARMA**(*p*, *q*) model, we need to rely on the difference equation to get the weights

$$(1 - \phi_1 z - \phi_2 z^2 - \cdots)(\psi_0 + \psi_1 z + \psi_2 z^2 + \cdots) = (1 + \theta_1 z + \theta_2 z^2 + \cdots).$$

The first few values are

$$\begin{aligned} \psi_0 &= 1 \\ \psi_1 - \phi_1\psi_0 &= \theta_1 \\ \psi_2 - \phi_1\psi_1 - \phi_2\psi_0 &= \theta_2 \\ \psi_3 - \phi_1\psi_2 - \phi_2\psi_1 - \phi_3\psi_0 &= \theta_3 \\ &\vdots \end{aligned}$$

✓The ψ-weights satisfy the homogeneous difference equation given by

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

Outline of ARIMA models

- **Basis of autoregressive moving average models**
- **Difference equations**
- **Autocorrelation (ACF) and partial autocorrelation (PACF)**
- **Forecasting**
- Estimation
- Integrated Models for Nonstationary Data
- Multiplicative Seasonal ARIMA Models

ACF and PACF: ACF of a MA model

ACF of a MA(q) model $X_t = \theta(B)W_t$, where $\theta(B) = 1 + \theta_1 B + \theta_2 B^2 + \cdots + \theta_q B^q$

- Since the **MA**(q) is a stationary process of second order, then

$$E(x_t) = \sum_{j=0}^{q} \theta_j E(w_{t-j}) = 0,$$

$$\gamma(h) = \operatorname{cov}(x_{t+h}, x_t) = \operatorname{cov}\left(\sum_{j=0}^{q} \theta_j w_{t+h-j}, \sum_{k=0}^{q} \theta_k w_{t-k} \right)$$

$$= \begin{cases} \sigma_w^2 \sum_{j=0}^{q-h} \theta_j \theta_{j+h}, & 0 \le h \le q \\ \boxed{0 \qquad\qquad h > q.} \end{cases}$$

Note that the ACF is symmetric. It is only needed to show ACF for h>=0

ACF: ACF of a general ARMA

Let a general **ARMA**(p, q): $\phi(B)X_t = \theta(B)W_t$, we write the process as

- ✓$X_t = \sum_{j=0}^{\infty} \psi_j W_{t-j}$,
- ✓where the ψ –weights can be estimated through difference equations
- ✓The mean $E[X_t] = 0$
- ✓The covariance function: $\gamma(h) = cov(X_{t+h}, X_t) = \sigma_W^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h}$
- ✓NB: two problems, 1) first to get ψ –weights, 2) ***there are infinite term terms***

An alternative way to write the covariance from the original definition $\phi(B)X_t = \theta(B)W_t$

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

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

PACF: Partial AutoCorrelation Function

- For **MA**(*q*) models, the ACF is zero for lags greater than *q*.
  - The ACF values --> *q* in **MA**(*q*), i.e., (the order of the dependence)
  - But ACF alone tells us little about *p* in **AR**(*p*) or **ARMA**(*p, q*)
  - For AR models ← the *partial autocorrelation function (PACF)*

- The PACF is defined as the partial correlation between X_{t+h}, X_t given Z={x_{t+1}, $x_{t+2}, \ldots, x_{t+h-1}$}
  - ✓ $\rho_{h|Z} = corr(X_{t+h} - \hat{X}_{t+h}, X_t - \hat{X}_t)$, with $\hat{X} = E[X_{t+h}|Z]$,$\hat{X}_t = E[X_t|Z]$ from regression
  - ✓The correlation between X_{t+h} and X_t is taking away the linear effect of *Z* from them
  - ✓Therefore, it is a two-step process to estimate PACF,
    - 1) regression mean
    - 2) estimate ρ for X-mean

PACF: definition

The PACF for the random process X_t, $\phi_{hh} = corr(X_{t+h}, X_t | x_{t+h-1}, \ldots, x_{t+1})$

- ✓ $\hat{X}_{t+h} = \beta_1 X_{t+h-1} + \beta_2 X_{t+h-2} + \cdots + \beta_{h-1} X_{t+1}$
- ✓ $\hat{X}_t = \beta_1 X_{t+h-1} + \beta_2 X_{t+h-2} + \cdots + \beta_{h-1} X_{t+1}$
- ✓ Because of the **ARMA** stationarity (mean and variance of X_t and X_{t+h} do not change), the coefficients $\beta_1, \beta_2, \ldots, \beta_{h-1}$ of both regression are the same

Therefore, the PACF of a stationary process X_t, denoted by ϕ_{hh} for $h = 1,2, \ldots$, can be defined as:

- ✓ $\phi_{11} = corr(X_{t+1}, X_t) = \rho(1)$
- ✓ $\phi_{hh} = corr\left(X_{t+h} - \hat{X}_{t+h}, X_t - \hat{X}_t\right)$, for $h \geq 2$
- ✓ NB: for time lag *h* > **AR**(*p*) order *p*, the PACF=0

PACF: examples (1)

Example (1): PACF of AR(1): $X_t = \phi X_{t-1} + W_t$ with $|\phi| \leq 1$

- ✓$\phi_{11} = corr(X_{t+1}, X_t) = \rho(1) = \phi$
- ✓To calculate ϕ_{22}, we need to first estimate the $\hat{X}_{t+2}$ and $\hat{X}_t$ as follows:
- ✓$E\left[(X_{t+2} - \hat{X}_{t+2})^2\right] = E[(X_{t+2} - \beta X_{t+1})^2] = \gamma(0) + 2\beta\gamma(1) + \beta^2\gamma(0)$ (regression: to minimize value)

  ➔ $\beta = \frac{\gamma(1)}{\gamma(0)} = \rho(1) = \phi$ (This will minimize the above cost function)

- ✓For X_t, $E\left[(X_t - \hat{X}_t)^2\right] = E[(X_t - \beta X_{t+1})^2] = \gamma(0) + 2\beta\gamma(1) + \beta^2\gamma(0)$ (similar as above)

  ➔ $\beta = \frac{\gamma(1)}{\gamma(0)} = \rho(1) = \phi$

  ➔ $\phi_{22} = corr\left(X_{t+2} - \hat{X}_{t+2}, X_t - \hat{X}_t\right) = corr(X_{t+2} - \phi X_{t+1}, X_t - \phi X_{t+1})$

  ➔ $\phi_{22} = corr(W_{t+2}, X_t - \phi X_{t+1}) = 0$ (NB: for lag > order p=1 here, the PACF=0)

PACF: examples (2)

Example (2): PACF of AR(p)

- $X_{t+h} = \sum_{j=1}^{p} \phi_j X_{t+h-j} + W_{t+h}$ with $|\text{Roots of } \phi(z)| > 1$ (causal stationary).
- For *h* > *p*, the regression of X_{t+h} in terms of $\{X_{t+1,} X_{t+2,} \dots, X_{t+h-1}\}$ as

  $$\hat{X}_{t+h} = \sum_{j=1}^{p} \phi_j X_{t+h-j}$$

- Based on above equation, after lag *h* > *p*, the time series are not correlated.

  ✓For *h* > *p*, we have $\phi_{hh} = corr\left(X_{t+h} - \hat{X}_{t+h}, X_t - \hat{X}_t\right) = corr\left(W_{t+h}, X_t - \hat{X}_t\right) = 0$

  ✓For $h \leq p$, we have $\phi_{hh} = \phi_h$

Example (3): PACF of MA(q)

- For invertible MA(q), $X_t = -\sum_{j=1}^{\infty} \pi_j X_{t-j} + W_t$
- Lets take MA(1) to look properties of PACF
- $\phi_{hh} = \frac{(-\theta)^h (1-\theta^2)}{1-\theta^{2(h+1)}}$ (not zero after q=1)

	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

Outline of ARIMA models

- Basis of autoregressive moving average models
- Difference equations
- Autocorrelation (ACF) and partial autocorrelation (PACF)
- **Forecasting**
- Estimation
- Integrated Models for Nonstationary Data
- Multiplicative Seasonal ARIMA Models

ARMA: forecasting ARMA process

- For an established causal and invertible **ARMA**(*p*, *q*) for predicting expected values (variances) for upcoming times: $X_{n+m}, m = 1,2, \ldots$
  - ✓ $\phi(B)X_t = \theta(B)W_t, W_t \sim iid\ N(0, \sigma_w^2),$
  - ✓ based on data collected at present, $X_{1:n} = \{X_1, \ldots, X_n\}$
- The prediction is usually represented by
  - ✓ $X_{n+m}^n = \mathrm{E}[X_{n+m}|X_n]$ as the format of $X_{n+m}^n = \alpha_0 + \sum_{k=1}^n \alpha_k X_k$
- Best Linear Prediction (**BLP**) for Stationary Process
  - ✓ The BLP is to estimate $\alpha_0, \alpha_1, \ldots, \alpha_k$ by solving:
  - ✓ $E[(X_{n+m} - X_{n+m}^n)X_k] = 0$ for $k = 0,1, \ldots, n$

so the *mean-square prediction error* can be written as

$$P_{n+m}^n = \mathrm{E}(x_{n+m} - \tilde{x}_{n+m})^2 = \sigma_w^2 \sum_{j=0}^{m-1} \psi_j^2 .$$

ARMA: forecasting (one-step-ahead)

- Let look at ***one-step-ahead prediction***, the BLP of X_{n+1}^n is

$$X_{n+1}^n = \phi_{n1}X_n + \phi_{n2}X_{n-1} + \cdots + \phi_{nn}X_1$$

✓Let $\phi_n = (\phi_{n1}, \phi_{n2}, \dots, \phi_{nn})^T$, $\gamma_n = (\gamma(1), \gamma(2), \dots, \gamma(n))^T$

✓Then, $\Gamma_n = \{\gamma(k-j)\}_{j,k=1}^n$ (a n*n matrix)

✓Then the coefficients: $\phi_n = \boldsymbol{\Gamma_n^{-1}}\gamma_n$

✓The mean squared errors is: $P_{n+1}^n = E[(X_{n+1} - X_{n+1}^n)^2] = \gamma_n' \boldsymbol{\Gamma_n^{-1}}\gamma_n$

ARMA: forecasting (m-step-ahead)

- The BLP of X_{n+m} for any $m \geq 1$, i.e., multi-step ahead prediction. Then the model becomes

$$\boxed{x_{n+m}^{n} = \phi_{n1}^{(m)} x_n + \phi_{n2}^{(m)} x_{n-1} + \cdots + \phi_{nn}^{(m)} x_1}$$

where $\{\phi_{n1}^{(m)}, \phi_{n2}^{(m)}, \ldots, \phi_{nn}^{(m)}\}$ satisfy the prediction equations,

$$\sum_{j=1}^{n} \phi_{nj}^{(m)} \gamma(k-j) = \gamma(m+k-1), \quad k = 1, \ldots, n.$$

$\gamma_n^{(m)} = (\gamma(m), \ldots, \gamma(m+n-1))'$, and $\phi_n^{(m)} = (\phi_{n1}^{(m)}, \ldots, \phi_{nn}^{(m)})'$ are $n \times 1$ vectors.

$$\Gamma_n \phi_n^{(m)} = \gamma_n^{(m)} \qquad P_{n+m}^{n} = \mathrm{E}\left(x_{n+m} - x_{n+m}^{n}\right)^2 = \gamma(0) - \gamma_n^{(m)\prime} \Gamma_n^{-1} \gamma_n^{(m)}.$$

Another useful algorithm for calculating forecasts was given by Brockwell and Davis (1991).

ARMA: forecasting ARMA example

Forecasting the Recruitment Series

$$x_{n+m}^n = 6.74 + 1.35x_{n+m-1}^n - .46x_{n+m-2}^n$$

$$x_t^s = x_t \text{ when } t \le s \quad \text{for } n = 453 \text{ and } m = 1, 2, \ldots, 12.$$

Recall that $\hat{\sigma}_w^2 = 89.72$,

$$\boxed{\begin{aligned} P_{n+1}^n &= 89.72, \\ P_{n+2}^n &= 89.72(1 + 1.35^2), \\ P_{n+3}^n &= 89.72(1 + 1.35^2 + [1.35^2 - .46]^2), \end{aligned}}$$

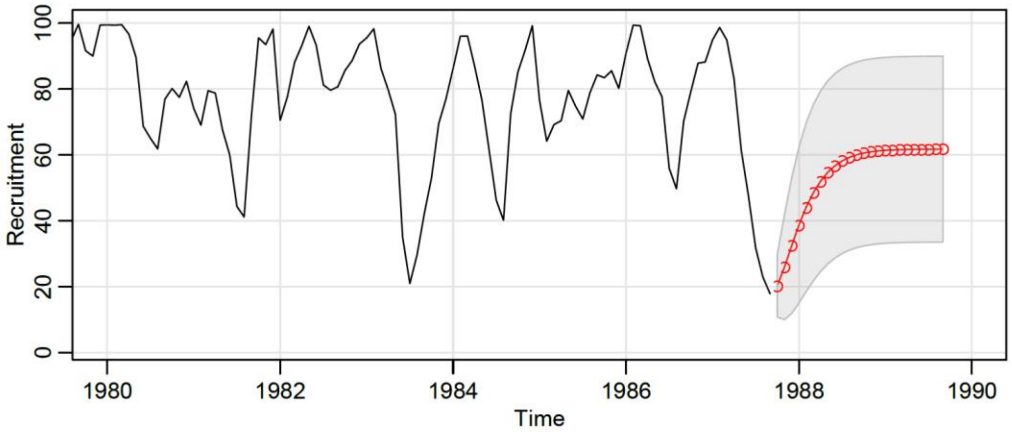


Fig. 3.7. Twenty-four month forecasts for the Recruitment series. The actual data shown are from about January 1980 to September 1987, and then the forecasts plus and minus one standard error are displayed.

AVANCEZ

1829

CHALMERS

UNIVERSITY OF TECHNOLOGY