Unit 11: ARMA Autocorrelation and Partial Autocorrelation Functions

Jeffrey Woo

Department of Statistics, University of Virginia

Spring 2020

Readings for Unit 11

Textbook chapter 3.3.

Last Unit

1. ARMA(p,q)
2. Condition for causality
3. Condition for invertibility

This Unit

1. ACF for MA(q)
2. ACF for Causal ARMA(p,q)
3. Partial Autocorrelation Function (PACF)

Motivation

In this unit we will study the autocorrelation and partial autocorrelation functions for ARMA processes.

MA(q) Process

Let's start with an MA(q) process

$$x_t = w_t + \theta_1 w_{t-1} + \theta_2 w_{t-2} + \cdots + \theta_q w_{t-q} = \sum_{j=0}^{q} \theta_j w_{t-j},$$

where we have written $\theta_0 = 1$. Then

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

Autocovariance for MA(q)

The autocovariance function is

$$\begin{aligned}\gamma(h) = cov(x_t, x_{t+h}) &= E\Big[\sum_{j=0}^{q} \theta_j w_{t-j} \sum_{j'=0}^{q} \theta_{j'} w_{t+h-j'}\Big] \\ &= \sum_{j=0}^{q}\sum_{j'=0}^{q} \theta_j \theta_{j'} E(w_{t-j} w_{t+h-j'}).\end{aligned}$$

Autocovariance for MA(q)

Recall that $E(w_s w_t) = \sigma_w^2$ if $s = t$ and $E(w_s w_t) = 0$ otherwise. So we have

$$\gamma(h) = \begin{cases} \sigma_w^2 \sum_{j=0}^{q-h} \theta_j \theta_{j+h}, & 0 \leq h \leq q, \\ 0, & h \geq q+1. \end{cases} \tag{1}$$

ACF for MA(q)

Recall that $\gamma(h) = \gamma(-h)$, so we will only need the values for $h \geq 0$. Dividing $\gamma(h)$ by $\gamma(0)$ in (1), we obtain the autocorrelation function (ACF) of an MA(q) model

$$\rho(h) = \begin{cases} \frac{\sum_{j=0}^{q-h} \theta_j \theta_{j+h}}{1+\theta_1^2+\cdots+\theta_q^2}, & 0 \leq h \leq q, \\ 0, & h \geq q+1. \end{cases} \tag{2}$$

ACF for Causal ARMA(p,q)

We have seen in (2), for MA(q) models, the ACF will be zero for lags greater than q. Moreover, because $\theta_q \neq 0$, $\rho(q) = \theta_0\theta_q/(1+\theta_1^2+\cdots+\theta_q^2) \neq 0$. Thus, the ACF provides information about the order of the dependence for a MA model. How about ARMA or AR models?

Causal ARMA(p,q)

Now we discuss causal ARMA(p, q) model

$$\phi(B)x_t = \theta(B)w_t,$$

where the roots of $\phi(z)$ are outside the unit circle. Since the model is causal, we have the MA(∞) representation

$$x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j}, \quad \text{where} \quad \psi(z) = \frac{\theta(z)}{\phi(z)} = \sum_{j=0}^{\infty} \psi_j z^j. \tag{3}$$

Autocovariance for Causal ARMA(p,q)

It follows that $E(x_t) = 0$ and by (1), the autocovariance function of x_t is given by

$$\gamma(h) = cov(x_t, x_{t+h}) = \sigma_w^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h}, \quad h \geq 0.$$

Autocovariance for Causal AR(1)

To motivate the idea, consider a causal AR(1) model $x_t = \phi_1 x_{t-1} + w_t$. Then

Autocovariance for Causal AR(1)

Here, we have used causality: x_{t-2} only depends on $w_{t-2}, w_{t-3}, \ldots$ and hence is independent of w_{t-1} and w_t. Consequently, $cov(w_{t-1}, x_{t-2}) = 0$ and $cov(w_t, x_{t-2}) = 0$. The ACF for an AR(1) model is very different from an MA(1) model whose ACF is zero at lag 2.

Autocovariance for Causal AR(1)

Question: Why is $\gamma(2) \neq 0$ for an AR(1) process?

Removing Connection

Consider the following,

because of causality. Thus, after removing the connection x_{t-1}, the covariance between x_t and x_{t-2} is zero.

Notation

One way to remove linear connections is through **linear regression**. Let $\hat{x}_{t+h}$ denote the regression of x_{t+h} on $\{x_{t+h-1}, x_{t+h-2}, \ldots, x_{t+1}\}$, which we write as

$$\hat{x}_{t+h} = \beta_1 x_{t+h-1} + \beta_2 x_{t+h-2} + \cdots + \beta_{h-1} x_{t+1}. \tag{4}$$

Here we do not include the intercept assuming the mean of x_t is zero. Otherwise, replace x_t with $x_t - \mu_x$.

Notation

In addition, let $\hat{x}_t$ denote the regression of x_t on $\{x_{t+1}, x_{t+2}, \ldots, x_{t+h-1}\}$, then

$$\hat{x}_t = \beta_1 x_{t+1} + \beta_2 x_{t+2} + \cdots + \beta_{h-1} x_{t+h-1}. \tag{5}$$

Partial Autocorrelation Function

The **partial autocorrelation function (PACF)** of a stationary process x_t, denoted by ϕ_{hh}, for $h = 1, 2, \ldots$, is

$$\phi_{11} = corr(x_{t+1}, x_t) = \rho(1) \tag{6}$$

and

$$\phi_{hh} = corr(x_{t+h} - \hat{x}_{t+h}, x_t - \hat{x}_t), \quad h \geq 2. \tag{7}$$

Note that, the 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.

Partial Autocorrelation Function

Now we consider PACF of a causal AR(1) model: $x_t = \phi x_{t-1} + w_t$, with $|\phi| < 1$. By definition, $\phi_{11} = corr(x_1, x_0) = \rho(1) = \phi$. To calculate ϕ_{22}, consider the regression of x_{t+2} on x_{t+1}, say $\hat{x}_{t+2} = \beta x_{t+1}$.

Partial Autocorrelation Function

Recall in linear regression we seek to minimize the error sum of squares (SSE). In this setting, we seek to minimize

Partial Autocorrelation Function

Next, consider the regression of x_t on x_{t+1}, say $\hat{x}_t = \beta x_{t+1}$. We choose β to minimize

Partial Autocorrelation Function

In general, for a causal AR(p) model $x_h = \sum_{j=1}^{p} \phi_j x_{h-j} + w_h$. When $h > p$, the regression of x_h on $x_{h-1}, \ldots, x_1$ is

$$\hat{x}_h = \sum_{j=1}^{p} \phi_j x_{h-j}.$$

Partial Autocorrelation Function

Thus, when $h > p$, by causality,

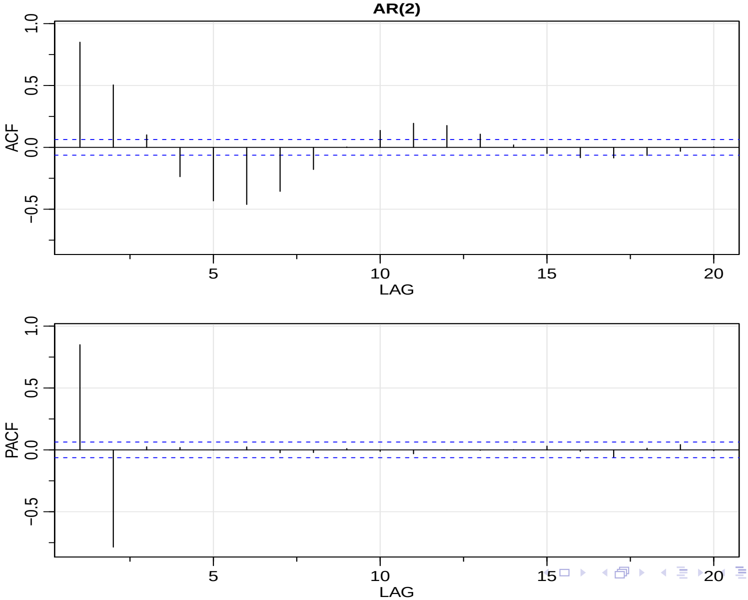
$$\phi_{hh} = corr(x_h - \hat{x}_h, x_0 - \hat{x}_0) = corr(w_h, x_0 - \hat{x}_0) = 0.$$

Summary

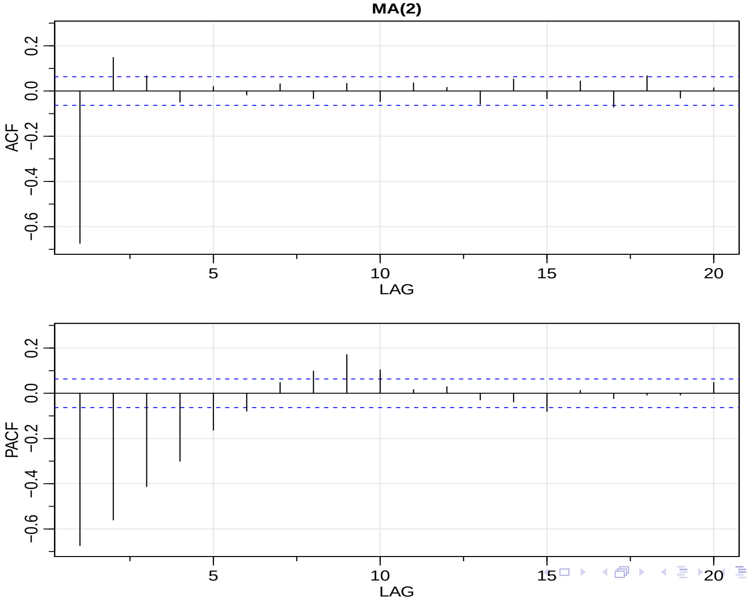
- The ACF of MA(q) model cuts at lag q. The PACF of an AR(p) model cuts at lag p.
- Identification of an MA(q) model is best done with ACF; identification of an AR(p) model is best done with PACF.
- The PACF between x_t and x_{t-h} is the conditional correlation between x_t and x_{t-h}. It is conditional on $x_{t-h+1}, \cdots, x_{t-1}$, the set of observations that come **between time** t **and** $t-h$.

1. ACF for MA(q) Processes
2. ACF for Causal ARMA(p,q) Processes
3. Partial Autocorrelation Function
4. Worked Examples

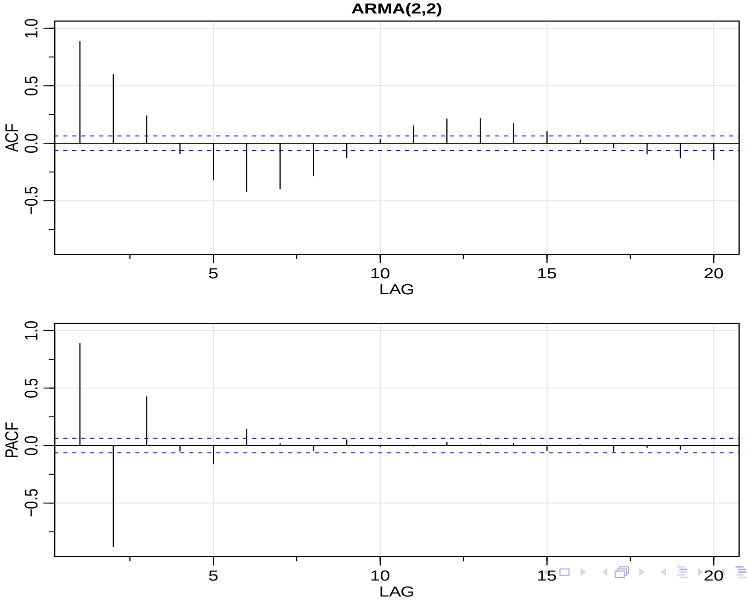
ACF and PACF of Causal AR(2)



ACF and PACF of Invertible MA(2)



ACF and PACF of Causal and Invertible ARMA(2,2)



ACF and PACF of Causal AR and Invertible MA

(From page 99, Table 3.1 of text)

	AR(p)	**MA(q)**	**ARMA(p,q)**
ACF	Decay	0 after lag q	Decay
PACF	0 after lag p	Decay	Decay

Fish Population Example

This time series from "recruit.dat" contains data on fish population in the central Pacific Ocean. The numbers represent the number of new fish in the years 1950-1987. **Question**: Based on the ACF and PACF plots, what process do you think is most likely to describe this time series?

Fish Population Example