# 3.2 Autocorrelation and Partial Autocorrelation

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

We begin by exhibiting the ACF of an MA($q$) process, $x_t = \theta(B)w_t$, where $\theta(B) = 1 + \theta_1 B + \ldots + \theta_q B^q$. Because $x_t$ is a finite linear combination of white noise terms, the process is stationary with mean

using *

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

where we have written $\theta_0 = 1$, and with autocovariance function

$$\gamma(h) = Cov(x_{t+h}, x_t) = \begin{cases} \sigma_w^2 \sum_{j=0}^{q-h} \theta_j \theta_{j+h}, & 0 \le h \le q \\ 0 & h > q. \end{cases} \tag{3.2.1}$$

Recall that $\gamma(h) = \gamma(-h)$, so we will only display the values for $h \ge 0$.

Note that $\gamma(q)$ cannot be zero because $\theta_q \neq 0$. The cutting off of $\gamma(h)$ after $q$ lags is the signature of the MA($q$) model.

Besides, the ACF of an MA($q$)is:

$$\rho(h) = = \begin{cases} \frac{\sum_{j=0}^{q-h} \theta_j \theta_{j+h}}{1+\theta_1^2+\ldots+\theta_q^2}, & 0 \leq h \leq q \\ 0 & h > q. \end{cases} \tag{3.2.2}$$

For a causal ARMA$(p, q)$ model, $\phi(B)x_t = \theta(B)w_t$, where the zeros of $\phi(z)$ are outside the unit circle, write

linear presentation for a causal ARMA(p,q)

$wn(0, \sigma_w^2)$

$$x_t = \sum_{j=0}^{\infty} \psi_j w_{t-j}. \tag{3.2.3}$$

It follows immediately that $E(x_t) = 0$ and the autocovariance function of $x_t$ is

$$\gamma(h) = Cov(x_{t+h}, x_t) = \sigma_w^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h}, \quad h \geq 0. \tag{3.2.4}$$

Besides, it is also possible to obtain a homogeneous difference equation directly in terms of $\gamma(h)$. First, we write

$$
\begin{aligned}
\gamma(h) = Cov(x_{t+h}, x_t) &= Cov\left(\sum_{j=1}^{p} \phi_j x_{t+h-j} + \sum_{j=0}^{q} \theta_j w_{t+h-j}, x_t\right) \\
&= \sum_{j=1}^{p} \phi_j \gamma(h-j) + \sigma_w^2 \sum_{j=h}^{q} \theta_j \psi_{j-h}, \qquad h \geq 0, \quad (3.2.5)
\end{aligned}
$$

where we have used the fact that, for $h \geq 0$,

$$
Cov(w_{t+h-j}, x_t) = Cov\left(w_{t+h-j}, \sum_{k=0}^{\infty} \psi_k w_{t-k}\right) = \psi_{j-h}\sigma_w^2.
$$

From (3.2.5), we can write a general homogeneous equation for the ACF of a causal ARMA process:

$$\gamma(h) - \phi_1\gamma(h-1) - \ldots - \phi_p\gamma(h-p) = 0, \qquad h \geq \max(p, q+1), \quad (3.2.6)$$

with initial conditions

$$\gamma(h) - \sum_{j=1}^{p} \phi_j\gamma(h-j) = \sigma_w^2 \sum_{j=h}^{q} \theta_j\psi_{j-h}, \qquad 0 \leq h < \max(p, q+1), \quad (3.2.7)$$

Dividing (3.2.6) and (3.2.7) through by $\gamma(0)$ will allow us to solve for the ACF, $\rho(h) = \gamma(h)/\gamma(0)$.

if AR(P) $\theta_j = 0$ $j = 1, \ldots, q$

**Example 3.5 (The ACF of an AR($p$))** *Consider AR($p$) process. To obtain its ACF, it follows immediately from* (3.2.6) *that*

$$\rho(h) - \phi_1\rho(h-1) - \ldots - \phi_p\rho(h-p) = 0, \qquad h \geq p. \tag{3.2.8}$$

*Let $z_1, \ldots, z_r$ denote the roots of $\phi(z)$, each with multiplicity $m_1, \ldots, m_r$, respectively, where $m_1 + \ldots + m_r = p$. Then, it can be shown that the general solution is*

$$\rho(h) = z_1^{-h}P_1(h) + z_2^{-h}P_2(h) + \ldots + z_r^{-h}P_r(h), \qquad h \geq p, \tag{3.2.9}$$

*where $P_j(h)$ is a polynomial in $h$ of degree $m_j - 1$.*

- For a causal model, all of the roots are outside the unit circle, $|z_i| > 1$, for $i = 1, \dots, r$.
- If all the roots are real, then $\rho(h)$ dampens exponentially fast to zero as $h \to \infty$.
- If some of the roots are complex, then they will be in conjugate pairs and $\rho(h)$ will dampen, in a sinusoidal fashion, exponentially fast to zero as $h \to \infty$.
- In the case of complex roots, the time series will appear to be cyclic in nature.
- This is also true for ARMA models in which the AR part has complex roots.

**Example 3.6 (The ACF of an ARMA(1, 1))** *Consider the ARMA(1, 1) process $x_t = \phi x_{t-1} + \theta w_{t-1} + w_t$, where $|\phi| < 1$. Based on* (3.2.6), *the autocovariance function satisfies*

causality

$\gamma(h) = \phi\gamma(h-1)$
$= \phi^2\gamma(h-2)$
$\vdots$
$= \phi^{h-1}\gamma(1)$

$$\gamma(h) - \phi\gamma(h-1) = 0, \qquad h = 2, 3, \ldots,$$

*and it can be shown that the general solution is*

$$\gamma(h) = \phi^{h-1}\gamma(1), \qquad h = 2, 3, \ldots \tag{3.2.10}$$

*To obtain the initial conditions, we use* (3.2.7)*:*

$$\gamma(0) = \phi\gamma(1) + \sigma_w^2[1 + \theta\phi + \theta^2]$$

*and*

$$\gamma(1) = \phi\gamma(0) + \sigma_w^2\theta.$$

*Solving for* $\gamma(0)$ *and* $\gamma(1)$, *we obtain:*

$$\gamma(0) = \sigma_w^2 \frac{1 + 2\theta\phi + \theta^2}{1 - \phi^2}$$

*and*

$$\gamma(1) = \sigma_w^2 \frac{(1 + \theta\phi)(\phi + \theta)}{1 - \phi^2}.$$

*Hence, the specific solution for* $h \geq 2$ *is*

$$\gamma(h) = \gamma(1)\phi^{h-1} = \sigma_w^2 \frac{(1 + \theta\phi)(\phi + \theta)}{1 - \phi^2} \phi^{h-1}$$

*Finally, dividing through by* $\gamma(0)$ *yields the ACF*

$$\rho(h) = \frac{(1 + \theta\phi)(\phi + \theta)}{1 - 2\phi\theta + \theta^2} \phi^{h-1}, \qquad h \geq 1 \tag{3.2.11}$$

decreasing exponentially as h increases

- The general pattern of $\rho(h)$ versus $h$ in (3.2.11) is not different from that of an AR(1) given in (3.1.8).
- It is unlikely that we will be able to tell the difference between an ARMA(1,1) and an AR(1) based solely on an ACF estimated from a sample.
- This consideration will lead us to the partial autocorrelation function.

### 3.2.1 The Partial Autocorrelation Function (PACF)

- We have seen in (3.2.2), for MA($q$) models, the ACF will be zero for lags greater than $q$.
- Because $\theta_q \neq 0$, the ACF will not be zero at lag $q$.
- The ACF provides a considerable amount of information about the order of the dependence when the process is a moving average process.
- If the process is ARMA or AR, the ACF alone tells us little about the orders of dependence.
- It is worthwhile pursuing a function that will behave like the ACF of MA models, but for AR models, namely, the partial autocorrelation function (PACF).

- Recall that if $X$, $Y$, and $Z$ are random variables, then the partial correlation between $X$ and $Y$ given $Z$ is obtained by regressing $X$ on $Z$ to obtain $\hat{X}$, regressing $Y$ on $Z$ to obtain $\hat{Y}$, and then calculating

$$\rho_{XY|Z} = Corr\{X - \hat{X}, Y - \hat{Y}\}.$$

- The idea is that $\rho_{XY|Z}$ measures the correlation between $X$ and $Y$ with the linear effect of $Z$ removed (or partialled out). If the variables are multivariate normal, then this definition coincides with $\rho_{XY|Z} = Corr(X, Y|Z)$.

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

$\phi x_{t-2} + w_{t-1}$ $\quad$ $x_{t-2} = \sum_{j=0}^{\infty} \psi_j w_{t-2-j}$

$$
\begin{aligned}
\gamma_x(2) = Cov(x_t, x_{t-2}) &= Cov(\phi x_{t-1} + w_t, x_{t-2}) \\
&= Cov(\phi^2 x_{t-2} + \phi w_{t-1} + w_t, x_{t-2}) = \phi^2 \gamma(0).
\end{aligned}
$$

$= \phi^2 Var(x_{t-2}) + \phi\, \cancel{Cov(w_{t-1}, x_{t-2})} + \cancel{Cov(w_t, x_{t-2})}$

- This result follows from causality because $x_{t-2}$ involves $\{w_{t-2}, w_{t-3}, \ldots\}$, which are all uncorrelated with $w_t$ and $w_{t-1}$.

- The correlation between $x_t$ and $x_{t-2}$ is not zero, as it would be for an MA(1), because $x_t$ is dependent on $x_{t-2}$ through $x_{t-1}$.

- Suppose we break this chain of dependence by removing (or partial out) the effect $x_{t-1}$. That is, we consider the correlation between $x_t - \phi x_{t-1}$ and $x_{t-2} - \phi x_{t-1}$, because it is the correlation between $x_t$ and $x_{t-2}$ with the linear dependence of each on $x_{t-1}$ removed.

In this way, we have broken the dependence chain between $x_t$ and $x_{t-2}$. In fact,

based on dep of AR(1) $x_t - \phi x_{t-1} = w_t$

$$Cov(x_t - \phi x_{t-1}, x_{t-2} - \phi x_{t-1}) = Cov(w_t, x_{t-2} - \phi x_{t-1}) = 0.$$

Hence, the tool we need is partial autocorrelation, which is the correlation between $x_s$ and $x_t$ with the linear effect of everything "in the middle" removed.

To formally define the PACF for mean-zero stationary time series, let $\hat{x}_{t+h}$, for $h \geq 2$, denote the regression of $x_{t+h}$ on $\{x_{t+h-1}, x_{t+h-2}, \ldots, x_{t+1}\}$, which we write as

$$\hat{x}_{t+h} = \beta_1 x_{t+h-1} + \beta_2 x_{t+h-2} + \ldots + \beta_{h-1} x_{t+1}. \tag{3.2.12}$$

No intercept term is needed in (3.2.12) because the mean of $x_t$ is zero (otherwise, replace $x_t$ by $x_t - \mu_x$ in this discussion).

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

$$\hat{x}_t = \beta_1 x_{t+1} + \beta_2 x_{t+2} + \ldots + \beta_{h-1} x_{t+h-1}. \tag{3.2.13}$$

Because of stationarity, the coefficients, $\beta_1, \ldots, \beta_{h-1}$ are the same in (3.2.12) and (3.2.13); we will explain this result in the next section, but it will be evident from the examples.

**Definition 3.9 (The partial autocorrelation function (PACF))** *The partial autocorrelation function of a stationary process, $x_t$, denoted by $\phi_{hh}$, for $h =$ $1, 2, \ldots,$ is*

$$\phi_{11} = Corr(x_{t+1}, x_t) = \rho(1), \tag{3.2.14}$$

*and*

$$\phi_{hh} = Corr(x_{t+h} - \hat{x}_{t+h}, x_t - \hat{x}_t), \qquad h \geq 2. \tag{3.2.15}$$

drop the effect of observations in the middle

$t \quad t{+}1 \quad \ldots \quad t{+}h{-}1 \quad t{+}h$

**Note 3.3**

- *The PACF, $\phi_{hh}$, is the correlation between $x_{t+h}$ and $x_t$ with the linear dependence of $\{x_{t+1}, \ldots, x_{t+h-1}\}$ on each, removed.*

- *If the process $x_t$ is Gaussian, then $\phi_{hh} = Corr(x_{t+h}, x_t | x_{t+1}, \ldots, x_{t+h-1})$; that is, $\phi_{hh}$ is the correlation coefficient between $x_{t+h}$ and $x_t$ in the bivariate distribution of $(x_{t+h}, x_t)$ conditional on $\{x_{t+1}, \ldots, x_{t+h-1}\}$.*

**Example 3.7 (The PACF of an AR(1))** *Consider the PACF of the AR(1) process given by $x_t = \phi x_{t-1} + w_t$, with $|\phi| < 1$. By definition, $\phi_{11} = \rho(1) = \phi$. To calculate $\phi_{22}$, consider the regression of $x_{t+2}$ on $x_{t+1}$, say, $\hat{x}_{t+2} = \beta x_{t+1}$. We choose $\beta$ to minimize*

t t+1 t+2

$$E(x_{t+2} - \hat{x}_{t+2})^2 = E(x_{t+2} - \beta x_{t+1})^2 = \gamma(0) - 2\beta\gamma(1) + \beta^2\gamma(0). \Rightarrow \beta = \frac{\gamma(1)}{\gamma(0)}$$

$=0$

*Taking derivatives with respect to $\beta$ and setting the result equal to zero, we have $\beta = \gamma(1)/\gamma(0) = \rho(1) = \phi$. Next, consider the regression of $x_t$ on $x_{t+1}$, say $\hat{x}_t = \beta x_{t+1}$. We choose $\beta$ to minimize*

$$E(x_t - \hat{x}_t)^2 = E(x_t - \beta x_{t+1})^2 = \gamma(0) - 2\beta\gamma(1) + \beta^2\gamma(0).$$

*This is the same equation as before, so $\beta = \phi$. Hence,*

PACF

$$\begin{aligned}\phi_{22} &= Corr(x_{t+2} - \hat{x}_{t+2}, x_t - \hat{x}_t) = Corr(\underbrace{x_{t+2} - \phi x_{t+1}}_{AR(1)}, x_t - \phi x_{t+1}) \\ &= Corr(w_{t+2}, x_t - \phi x_{t+1}) = 0\end{aligned}$$

t+2 ) t, t+1

*by causality. Thus,* $\phi_{22} = 0$. *In the next example, we will see that in this case,* $\phi_{hh} = 0$ *for all* $h > 1$.

AR(1) $\phi_{11} = \phi$

$\phi_{22} = \phi_{33} = 0$

**Example 3.8 (The PACF of an AR($p$))** *The model implies* $x_{t+h} = \sum_{j=1}^{p} \phi_j x_{t+h-j} + w_{t+h}$, *where the roots of* $\phi(z)$ *are outside the unit circle. When* $h > p$, *the regression of* $x_{t+h}$ *on* $\{x_{t+1}, \ldots, x_{t+h-1}\}$, *is*

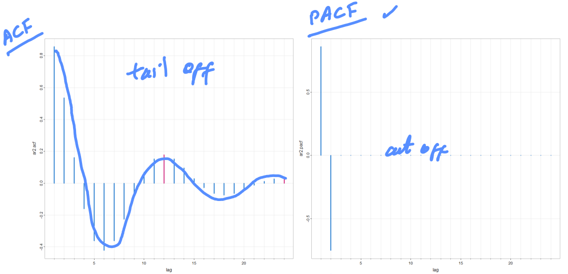
$$\hat{x}_{t+h} = \sum_{j=1}^{p} \phi_j x_{t+h-j}$$

*Thus, when* $h > p$,

$$\phi_{hh} = Corr(x_{t+h} - \hat{x}_{t+h}, x_t - \hat{x}_t) = Corr(w_{t+h}, x_t - \hat{x}_t) = 0, \qquad h > p$$

*because, by causality,* $x_t - \hat{x}_t$ *depends only on* $\{w_{t+h-1}, w_{t+h-2}, \ldots\}$. *When* $h \leq p$, $\phi_{pp}$ *is not zero, and* $\phi_{11}, \ldots, \phi_{p-1,p-1}$ *are not necessarily zero. We will see later that, in fact,* $\phi_{pp} = \phi_p$.

*Figure 3.3 shows the ACF and the PACF of the AR(2) with* $\phi_1 = 1.5$ *and* $\phi_2 = -.75$.




MA(-q) ACF cuts off at lag q PACF tails off (decreases gradually)

Figure 3.3: The ACF and PACF of an AR(2) model with $\phi_1 = 1.5$ and $\phi_2 = -.75$

**Example 3.9 (The PACF of an Invertible MA($q$))** *For an invertible MA($q$), we can write* $x_t = \sum_{j=1}^{\infty} \pi_j x_{t-j} + w_t$ *.Moreover, no finite representation exists. From this result, it should be apparent that the PACF will never cut off, as in the case of an AR($p$). For an MA(1),* $x_t = w_t + \theta w_{t-1}$*, with* $|\theta| < 1$*, it can be shown that* $\phi_{22} = -\theta^2/(1+\theta^2+\theta^4)$*. For the MA(1) in general, we can show that*

$$\phi_{hh} = -\frac{(-\theta)^h(1-\theta^2)}{1-\theta^{2(h+1)}}, \qquad h \geq 1.$$

- The PACF for MA models behaves much like the ACF for AR models.
- The PACF for AR models behaves much like the ACF for MA models.
- Because an invertible ARMA model has an infinite AR representation, the PACF will not cut off.

| | 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 |

Table 3.1: Behavior of the ACF and PACF for ARMA models

**Example 3.10 (Preliminary Analysis of the Recruitment Series)** *We consider the problem of modeling the Recruitment series, discussed in previous chapters. There are 453 months of observed recruitment ranging over the years 1950-1987. The ACF and the PACF given in Figure 3.10 are consistent with the behavior of an AR(2). The ACF has cycles corresponding roughly to a 12-month period, and the PACF has large values for* $h = 1, 2$ *and then is essentially zero for higher order lags. Based on Table 3.1, these results suggest that a second-order* ($p = 2$) *autoregressive model might provide a good fit.*

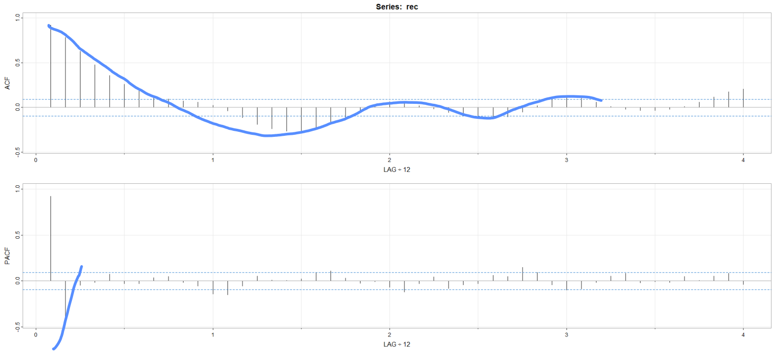
*We ran a regression using the data triplets*

$$\{(x; z_1, z_2) : \ (x_3; x_2, x_1), (x_4; x_3, x_2), \ldots, (x_{453}; x_{452}, x_{451})\}$$

*to fit a model of the form* $x_t = \phi_0 + \phi_1 x_{t-1} + \phi_2 x_{t-2} + w_t$ *for* $t = 3, 4, \ldots, 453$. *The estimates and standard errors (in parentheses) are* $\hat{\phi}_0 = 6.74(1.11)$, $\hat{\phi}_1 =$

1.35(.04), $\hat{\phi}_2 = -.46(.04)$, *and* $\sigma_w^2 = 89.72$. *The following R code can be used for this analysis. We use acf2 from astsa to print and plot the ACF and PACF.*

AR(2)



auto-covariance function of MA(q)

$\gamma(h) = Cov(x_{t+h}, x_t)$

$= Cov\left(\sum_{j=0}^{q} \theta_j w_{t+h-j}, \sum_{k=0}^{q} \theta_k w_{t-k}\right)$

$= \sum_{j=0}^{q} \sum_{k=0}^{q} \theta_j \theta_k \underbrace{Cov(w_{t+h-j}, w_{t-k})}$

$\neq 0$ if $t+h-j = t-k$

$\Rightarrow j = h+k$

$0 \leq \underbrace{j}_{h+k} \leq q$

$\Rightarrow k \leq q-h$

$= \sum_{k=0}^{q-h} \theta_{h+k} \theta_k \sigma_w^2$ if $0 \leq h \leq q$

slide 207 ARMA(p,q)

$x_t = \phi_1 x_{t-1} + \cdots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \cdots + \theta_q w_{t-q}$

$\gamma(h) = Cov(x_{t+h}, x_t)$

$= Cov\left(\sum_{j=1}^{p} \phi_j x_{t+h-j} + \sum_{k=0}^{q} \theta_k w_{t+h-k}, x_t\right)$

$$= \sum_{j=1}^{p} \phi_j \underbrace{Cov(x_{t+h-j}, x_t)}_{\gamma(h-j)}$$

$$+ \sum_{k=0}^{q} \theta_k \underbrace{Cov(w_{t+h-k}, x_t)}_{**}$$

$$= \sum_{j=1}^{p} \phi_j \gamma(h-j) + \sum_{k=h}^{q} \psi_{k-h} \sigma_w^2$$

$$Cov(w_{t+h-k}, x_t) = Cov\left(w_{t+h-k}, \sum_{j=0}^{\infty} \psi_j w_{t-j}\right)$$

linear presentation

$$= \sum_{j=0}^{\infty} \psi_j \underbrace{Cov(w_{t+h-k}, w_{t-j})}_{\neq 0 \text{ if } \cancel{t}+h-k=\cancel{t}-j}$$

$j = k-h$

$$= \psi_{k-h} \sigma_w^2$$

---

Slide 208 general homogeneous equation of ACF

$$\gamma(h) - \phi_1 \gamma(h-1) - \dots - \phi_p \gamma(h-p) = 0 \qquad h \geq \max(p, q+1)$$

initial condition: $\gamma(h) - \sum_{j=1}^{p} \phi_j \gamma(h-j) = \sigma_w^2 \sum_{j=h}^{q} \theta_j \psi_{j-h}$

if $0 \leq h < \underbrace{\max(p, q+1)}_{\text{why ?!}}$

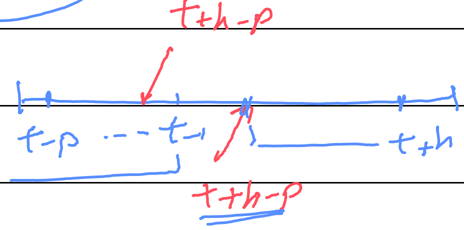
$$x_t = \phi_1 x_{t-1} + \dots + \phi_p x_{t-p} + w_t + \theta_1 w_{t-1} + \dots + \theta_q w_{t-q}$$

$h>0$

$$x_{t+h} = \phi_1 x_{t+h-1} + \dots + \phi_p x_{t+h-p} + w_{t+h} + \theta_1 w_{t+h-1} + \dots + \theta_q w_{t+h-q}$$

we have connection between $x_t$ & $x_{t+h}$ if

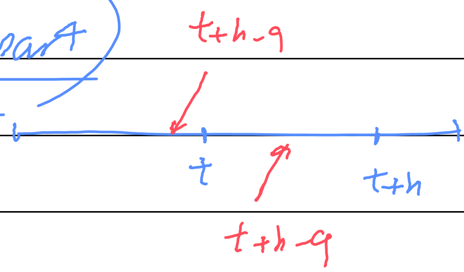
AR part




$t+h-p \leq t-1 \Rightarrow h \leq p-1$

OR

MA part




$t+h-q \leq t \Rightarrow h \leq q$

$$\left.\begin{array}{l} h \leq p-1 \\ h \leq q \end{array}\right\} \Rightarrow h \leq \max(p-1, q)$$

$$\Rightarrow h < \max(p, q+1)$$

Example 3.5 AR($p$), $p \geq 2$

general homogenous equation

$$\gamma(h) - \phi_1 \gamma(h-1) - \dots - \phi_p \gamma(h-p) = 0 \qquad h \geq \max(p, 2) = p$$

Slide 211 initial conditions to extract $\gamma(0)$ and $\gamma(1)$

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

$p=1$
$q=1$

$$\gamma(h) - \phi \gamma(h-1) = \sigma_w^2 \sum_{j=h}^{1} \theta_j \underbrace{\psi_{j-h}}_{\substack{\text{linear presentation} \\ \text{do coefficient matching}}}, \quad h=0,1$$

$\psi_0 = 1 \quad \psi_1 = \phi - \theta$

$\psi_2 = \phi^2 - \theta\phi$

if $h=0 \Rightarrow \gamma(0) - \phi\gamma(1) = \sigma_w^2 (\underbrace{\theta_0}_{1} \psi_0 + \theta_1 \psi_1)$

$\Rightarrow \gamma(0) = \sigma_w^2 (1 + \theta\phi - \theta^2) \quad (a)$

if $h=1 \Rightarrow \gamma(1) - \phi\gamma(0) = \sigma_w^2 (\theta_1 \psi_0)$

$\Rightarrow \gamma(1) = \theta\sigma_w^2 + \phi\gamma(0) \quad (b)$

solve (a) and (b) to get $\gamma(0), \gamma(1)$

Slide 220

$$E(x_{t+2} - \hat{x}_{t+2})^2 = E(x_{t+2} - \beta x_{t+1})^2$$

assume that $E(x_{t+2}) = 0$ otherwise subtract mean

$$= E(x_{t+2}^2) - \beta E(x_{t+2} x_{t+1}) + \beta^2 E(x_{t+1}^2)$$

$$= Var(x_{t+2}) - \beta\, Cov(x_{t+2}, x_{t+1}) + \beta^2 Var(x_{t+1})$$

$$= \gamma(0) - \beta\gamma(1) + \beta^2\gamma(0)$$