Lecture 4: Basic Time Series Models (Continued)

Walid Sharabati

Purdue University

Latest Update
September 5, 2016
For the stationary stochastic process $X(t)$ or $X_t$ we have

$$\rho_\tau = \frac{\gamma_\tau}{\gamma_0} = \frac{\gamma_\tau}{\sigma^2}.$$ 

1. $\rho_0 = 1$.

2. Covariance is symmetric, $\rho_\tau = \rho_{-\tau}$.

$$\gamma_\tau = \text{cov}(X_t, X_{t+\tau}) = \text{cov}(X_t, X_{t-\tau}) = \text{cov}(X_{t-\tau}, X_t) = \gamma_{-\tau}.$$ 

Since $X_t$ is stationary.

3. $|\rho_\tau| \leq 1$.

4. A stochastic process $\Rightarrow$ unique ac.f. The converse is not necessarily true ($\not\Leftarrow$).
1 Purely random processes (iid noise): \( \{Z_t\} \sim IID(0, \sigma^2) \).

\[ \gamma_\tau = Cov(Z_t, Z_{t+\tau}) = \begin{cases} 
0, & \tau = \pm 1, \pm 2, \cdots \\
\sigma^2, & \tau = 0.
\end{cases} \]

Purely random processes are 2nd order stationary, mean and acv.f. don’t depend on \( t \).

\[ \rho_\tau = \frac{\gamma_\tau}{\gamma_0} = \begin{cases} 
1, & \tau = 0. \\
0, & \tau \neq 0.
\end{cases} \]

If the distribution is Gaussian then it’s strictly stationary.
2 Random walk (iid noise): \( \{ Z_t \} \sim IID(\mu, \sigma^2) \).

\[
X_0 = 0, \quad X_t = X_{t-1} + Z_t.
\]

\( X_0 = 0 \), so the process starts at 0 when \( t = 0 \).

\[
X_1 = X_0 + Z_1 = Z_1.
\]

\[
X_2 = X_1 + Z_2 = Z_1 + Z_2.
\]

\[
X_3 = X_2 + Z_3 = Z_1 + Z_2 + Z_3.
\]

\[ \vdots \]

\[
X_t = \sum_{i=1}^{t} Z_i.
\]
Example

share price on day $t = \text{share price on day } (t - 1) + \text{random noise}$.

Mean and Variance

$E(X_t) = \sum_{i=1}^{t} E(Z_i) = t \cdot \mu$, \quad $Var(X_t) = \sum_{i=1}^{t} Var(Z_i) = t \cdot \sigma^2$.

Mean and variance depend on $t$ so it is not a stationary time series.

$\nabla X_t = X_t - X_{t-1} = Z_t$. 
R code for a random walk

## Linear trend in time
data(rwalk) ## rwalk contains a simulated random walk
win.graph(width = 4.874, height = 2.5, pointsize = 8)
plot(rwalk, main = "", type='o', ylab='Random Walk', col='blue');
title(main = list("Time Series Plot of a Random Walk with the Regression Line", cex=1.3, col="red", font=2))

## Least squares regression (LSR) model for linear time trend
## Xt (rwalk) vs t
model1 = lm(rwalk~time(rwalk))
summary(model1) ## report the parameter estimates
anova(model1) ## display the anova table for model1

abline(model1) ## add the fitted least squares line from model1
Figure: Time series plot of a random walk with the regression line.
R code for a simulated random walk

## Simulated random walk or Brownian motion
set.seed(154) # so you can reproduce the results

# generate normally distributed random noise
w = rnorm(200,0,1)

# generate the time series, which is the sum of noise (w)
x = cumsum(w)

wd = w +.2
# generate the time series, which is the sum of noise (wd)
xd = cumsum(wd)

## Plot the time series
plot.ts(xd, ylim=c(-5,55), main="random walk")
lines(x)
lines(.2*1:200, lty="dashed")
Figure: Time series plot of a simulated random walk with the regression line.
3 Moving average processes $\text{MA}(q)$: $\{Z_t\} \sim IID(0, \sigma^2)$.

$$X_t = \beta_0 Z_t + \beta_1 Z_{t-1} + \beta_2 Z_{t-2} + \cdots + \beta_q Z_{t-q}.$$  

We may rescale $Z_t$ so that $\beta_0 = 1$.

Mean and Variance

$$E(X_t) = 0, \quad Var(X_t) = \sigma_Z^2 \sum_{i=0}^{q} \beta_i^2.$$
Example

MA(1): $X_t = Z_t + \beta_1 Z_{t-1}$. Find $\gamma_k, \rho_k$.

$$
\gamma_k = Cov(X_t, X_{t+k}) = \begin{cases} 
(1 + \beta_1^2) \sigma^2, & k = 0. \\
\beta_1 \sigma^2, & |k| = 1. \\
0, & \text{otherwise}.
\end{cases}
$$

$$
\rho_k = \frac{\gamma_k}{\gamma_0} = \frac{Cov(X_t, X_{t+k})}{Var(X_t)} = \begin{cases} 
\frac{1}{\beta_1}, & k = 0. \\
\frac{\beta_1}{1 + \beta_1^2}, & |k| = 1. \\
0, & |k| \neq 1.
\end{cases}
$$

- Since $E(X_t), \gamma_k, \text{and } \rho_k$ do not depend on $t$ (constants), the MA process is 2nd order stationary.
- If $\{Z_t\}$ are normally distributed so is $\{X_t\}$, which means that MA($q$) is a strictly stationary Gaussian process.
Exercise
For \( X_t = Z_t + \beta_1 Z_{t-1} + \beta_2 Z_{t-2} + \cdots + \beta_q Z_{t-q} \), show that:

\[
\gamma_k = \begin{cases} 
0, & k > q. \\
\sigma^2 \sum_{i=0}^{q-k} \beta_i \beta_{i+k}, & k = 0, 1, \cdots, q. \\
\gamma_{-k}, & k < 0.
\end{cases}
\]

\[
\rho_k = \begin{cases} 
0, & k > q. \\
1, & k = 0. \\
\sum_{i=0}^{q-k} \beta_i \beta_{i+k}, & k = 1, 2, \cdots, q. \\
\frac{\sum_{i=0}^{q} \beta_i^2}{\sum_{i=0}^{q} \beta_i^2}, & k = 0. \\
\rho_{-k} & k < 0.
\end{cases}
\]

- The ac.f. cuts off at lag \( q \), a feature/benchmark of MA(\( q \)) process.
data(ttrc)
sma <- SMA(ttrc[,"Close"], 20)
acf(tail(sma, 20), 9, main="", col ="red")
title(main=list("Correlogram of a Moving Average Process (r_k vs k)"))
Summary

1. Properties:
   1.1 $\rho_0 = 1$, $|\rho_\tau| \leq 1$.
   1.2 $\rho_\tau = \rho_{-\tau}$.
   1.3 A stochastic process $\Rightarrow$ unique ac.f. ($\not\Leftarrow$).

2. Purely random processes: $\{Z_t\} \sim IID(0, \sigma^2)$.
   $$\gamma_\tau = \begin{cases} 0, & |\tau| = 1, 2, \cdots, \\ \sigma^2, & \tau = 0. \end{cases} \quad \rho_\tau = \begin{cases} 1, & \tau = 0. \\ 0, & \tau \neq 0. \end{cases}$$

3. Random walk: $\{Z_t\} \sim IID(\mu, \sigma^2)$.
   $$X_0 = 0, \quad X_t = X_{t-1} + Z_t.$$  
   $$E(X_t) = t \cdot \mu, \quad Var(X_t) = t \cdot \sigma^2.$$  

4. Moving average MA($q$): $\{Z_t\} \sim IID(0, \sigma^2)$.
   $$X_t = Z_t + \beta_1 Z_{t-1} + \beta_2 Z_{t-2} + \cdots + \beta_q Z_{t-q}.$$  
   $$E(X_t) = 0, \quad Var(X_t) = \sigma_Z^2 \sum_{i=0}^{q} \beta_i^2.$$  


Benchmark Property

1. MA($q$): $\gamma_k = 0$ or $\rho_k = 0$ when $|k| > q$.
2. MA(1): $\gamma_k = 0$ for $|k| > 1$.
3. MA(2): $\gamma_k = 0$ for $|k| > 2$. 
R code to generate an MA(1) process

par(mfrow=c(2,1))

# in the expressions below, is a space and == is equal
plot(arima.sim(list(order=c(0,0,1), ma=.5), n=100), ylab="x"
     main=(expression(MA(1) theta==+.5)))

plot(arima.sim(list(order=c(0,0,1), ma=-.5), n=100), ylab="x"
     main=(expression(MA(1) theta==-.5)))
4 Autoregressive Process AR($p$).

Let $\{Z_t\} \overset{iid}{\sim} (0, \sigma^2_Z)$ be the noise. The autoregressive process with parameter $p$ is given by

$$X_t = \alpha_1 X_{t-1} + \cdots + \alpha_p X_{t-p} + Z_t.$$ 

AR($p$) is equivalent to multiple linear regression

$$Y = \beta_0 + \beta_1 X_1 + \cdots + \beta_{p-1} X_{p-1} + \epsilon.$$ 

Markov Process

$p = 1 \quad \rightarrow \quad \text{AR}(1) \quad \rightarrow \quad \text{Markov Chain Process}.$ 

$$X_t = \alpha_1 X_{t-1} + Z_t.$$
R code to generate an AR(1) process

par(mfrow=c(2,1))

# in the expressions below, is a space and == is equal
plot(arima.sim(list(order=c(1,0,0), ar=.9), n=100), ylab="x",
main=(expression(AR(1) phi==+.9)))

plot(arima.sim(list(order=c(1,0,0), ar=-.9), n=100), ylab="x",
main=(expression(AR(1) phi==-.9)))
**Back-Shift Operator** $B$

The backward shift operator $B$ is defined such that $BX_t = X_{t-1}$.

$$B^2X_t = B(BX_t) = BX_{t-1} = X_{t-2}.$$ 

In general,

$$B^jX_t = X_{t-j}, \quad \forall \ j.$$ 

**Example**

$$X_t = \alpha BX_t + Z_t \quad \Rightarrow \quad (1 - \alpha B)X_t = Z_t \quad \Rightarrow \quad X_t = \frac{Z_t}{1 - \alpha B}.$$ 

**Infinite Series Expansion**

$$\frac{1}{1 - t} = 1 + t + t^2 + \cdots = \sum_{i=0}^{\infty} t^i.$$ 

$$\frac{1}{1 - \alpha B} = 1 + \alpha B + \alpha^2 B^2 + \cdots = \sum_{i=0}^{\infty} \alpha^i B^i.$$
\[ \alpha B Z_t = \alpha Z_{t-1}. \]

If we apply the backshift operator on the moving average MA\((q)\) we get
\[
X_t = (1 + \alpha B + \alpha^2 B^2 + \cdots) Z_t.
\]
\[
= Z_t + \alpha Z_{t-1} + \alpha^2 Z_{t-2} + \cdots = \sum_{i=0}^{\infty} \alpha^i Z_{t-i}
\]

For the Markov process AR\((1)\) we have
\[
X_t = \alpha X_{t-1} + Z_t.
\]
\[
= \alpha(\alpha X_{t-2} + Z_{t-1}) + Z_t.
\]
\[
= \alpha^2 X_{t-2} + \alpha Z_{t-1} + Z_t.
\]
\[
= \alpha^2(\alpha X_{t-3} + Z_{t-2}) + \alpha Z_{t-1} + Z_t = \sum_{i=0}^{\infty} \alpha^i Z_{t-i}.
\]
• An AR(1) process $X_t$ may be expressed as an infinite-order MA process provided $|\alpha| < 1$ so that the sum converges.

• There is a duality between AR and MA processes.
  ○ AR processes may be written in MA form.

• Use backward shift operator instead of successive substitution.

$$AR(1) \rightarrow X_t = \alpha X_{t-1} + Z_t, \quad \text{But} \quad (1 - \alpha B)X_t = Z_t.$$ 

So,

$$X_t = \frac{Z_t}{1 - \alpha B} = (1 + \alpha B + \alpha^2 B^2 + \cdots)Z_t = Z_t + \alpha Z_{t-1} + \alpha^2 Z_{t-2} + \cdots.$$ 

Thus,

$$E(X_i) = 0, \quad Var(X_t) = (1 + \alpha^2 + \alpha^4 + \cdots)\sigma_Z^2 = \frac{1}{1 - \alpha^2} \cdot \sigma_Z^2.$$
**Autocovariance of the Autoregressive Process**

The autocovariance function is given by

$$\gamma_k = \text{cov} \left( X_t, X_{t+k} \right) = \text{cov} \left( \sum_{i=0}^{\infty} \alpha^i Z_{t-i}, \sum_{j=0}^{\infty} \alpha^j Z_{t+k-j} \right)$$

let $t - i = t + k - j \quad \rightarrow \quad j = i + k$.

$$= \text{cov} \left( \sum_{i=0}^{\infty} \alpha^i Z_{t-i}, \sum_{i=-k}^{\infty} \alpha^{i+k} Z_{t-i} \right) = \sum_{i=0}^{\infty} \alpha^i \alpha^{i+k} \text{cov} \left( Z_{t-i}, Z_{t-i} \right)$$

$$= \sum_{i=0}^{\infty} \alpha^{2i+k} \sigma_Z^2 = \alpha^k \left( \sum_{i=0}^{\infty} \alpha^2 \right)^i \sigma_Z^2 = \frac{\alpha^k}{1 - \alpha^2} \sigma_Z^2 = \alpha^k \cdot \sigma_{X_t}^2.$$
Autocorrelation of the Autoregressive Process

The autocorrelation function is given by

\[ \rho_k = \frac{\gamma_k}{\gamma_0} = \frac{\alpha^k}{1 - \alpha^2} = \alpha^k, \quad \text{provided } |\alpha| < 1. \]

- AR(1) is a second order stationary with
  \[ \rho_k = \alpha^{|k|}, \quad k = 0, \pm 1, \pm 2, \ldots. \]

Example

\[ X_t = 0.9X_{t-1} + Z_t, \quad \rho_k = 0.9^k. \]
\[ X_t = -0.9X_{t-1} + Z_t, \quad \rho_k = (-0.9)^k. \]
R code for the autocorrelation of an AR(2) process

```r
ar2.acf = ARMAacf(ar=c(1.5, -.75), ma=0, 24)[-1]
par(mfrow=c(1,2))

plot(ar2.acf, type="h", xlab="lag")
abline(h=0)
```
Invertible Processes (Duality Between AR & MA)

Definition
A process \( \{X_t\} \) or MA(q) is said to be invertible if the random distribution at time \( t \) (innovation) can be expressed as a convergent sum of present and past values of the process in the form

\[
Z_t = \sum_{j=0}^{\infty} \pi_j X_{t-j}, \quad \text{where} \quad \sum_{j=0}^{\infty} |\pi_j| < \infty.
\]

MA(\( \infty \)), this effectively means that the process can be rewritten in the form of an infinite order AR process whose coefficient form a convergent sum. with \( \sum_j |\pi_j| < \infty \), \( Z_t \) is an AR(\( \infty \)).

MA(q) : \[
X_t = \beta_0 Z_t + \beta_1 Z_{t-1} + \cdots + \beta_q Z_{t-q}
= (\beta_0 + \beta_1 B + \cdots + \beta_q B^q) Z_t = \theta(B) Z_t,
\]
where \( \theta(B) = \beta_0 + \beta_1 B + \cdots + \beta_q B^q \).
• MA($q$) process is invertible if the roots of the equation

$$\theta(B) = \beta_0 + \beta_1 B + \cdots + \beta_q B^q = 0$$

all lie outside the unit circle, where $B$ is regarded as a complex variable and not as an operator.

• This means the roots, which may be complex, have modulus greater than unity.

**Exercise**

MA(1): $X_t = Z_t + \beta Z_{t-1}$, find $\beta$ such that MA(1) is invertible.

**Solution:**

$$X_t = Z_t + \beta B Z_t = (1 + \beta B) Z_t$$

$|B| > 1 \Rightarrow |\beta| < 1$. Since $|\beta| < 1$, the process MA(1) is invertible.