Oct30 Notes

Hongdou Li

10/30/2018

Backshift Operator: B

$Bf(t) = f(t-1)$

$BY_t = Y_{t-1}$

$B^2Y_t == BBY_t = BY_{t-1} = Y_{t-2}$

$B^nY_t = Y_{t-n}$ for n=0,1,2,... * $B^0 = 1$

AR(p) Process

The time series $\{Y_t\}$ is called an **autoregressive process of order p** if

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \epsilon_t$$

where $\{\epsilon_t\} \sim WN(0, \sigma^2)$ and $\phi_1, \phi_2, ..., \phi_p$ are sonstants.

* An AR(p) process is **only stationary** if the **stationary condition** on the $\phi's$ is met.

We can rewrite this relationship using backshift operators in the following way:

$$Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - ... - \phi_p Y_{t-p} = \epsilon_t$$

$$B^0 Y_t - \phi_1 B^1 Y_t - \phi_2 B^2 Y_t - ...\phi_p B^p Y_t = \epsilon_t$$

$$(1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p) Y_t = \epsilon_t$$

$$(1 - \textstyle\sum_{i=1}^p \phi_i B^i) Y_t = \epsilon_t$$

we defile $\phi^p(z) = 1 - \sum_{i=1}^p \phi_i z^i$ to be the **generating function** of the AR(p) process. Using this, the AR(p) relationship can be written as :

$$\phi^p(B) Y_t = \epsilon_t$$

MA(q) process

A time series $\{Y_t\}$ is called a **moving average process of order q** if

$$Y_t = \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + ... + \theta_q \epsilon_{t-q}$$

where $\{\epsilon_t\} \sim WN(0, \sigma^2)$ and $\theta_!, \theta_2, ..., \theta_q$ are constants.

* $MA(q) = AR(\infty)$ (this is true for all q, we'll prove it for q = 1)

MA(1): $Y_t = \epsilon_t + \theta\epsilon_{t-1} => \epsilon_t = Y_t - \theta\epsilon_{t-1}$

$= \epsilon_t + \theta(Y_{t-1} - \theta\epsilon_{t-2})$

$= \epsilon_t + \theta Y_{t-1} - \theta^2 \epsilon_{t-2}$

$= \epsilon_t + \theta Y_{t-1} - \theta^2 (Y_{t-2} - \theta\epsilon_{t-3})$

$= \epsilon_t + \theta Y_{t-1} - \theta^2 Y_{t-2} + \theta^3 \epsilon_{t-3} \ldots\ldots$

that is $\phi_i = \theta^i(-1)^{i+1}$

This is True as longas **"intertibility conditions"** on the θ's are met.

We can rewrite the MA(q) relationship using backshift operator notation:

$$Y_t = B^0\epsilon_t + \theta_1 B\epsilon_t + \theta_2 B^2\epsilon_t + ... + \theta_q B^q\epsilon_t$$
$$Y_t = (1 + \theta_1 B + \theta_2 B^2 + ... + \theta_q B^q)\epsilon_t$$
$$Y_t = (1 + \sum_{j=1}^{q} \theta_j B^j)\epsilon_t$$

we define $\Theta^q(z) = 1 + \sum_{j=1}^{q} \theta_j z^j$ to be the **generating function** of the MA(q) process. Using this, the MA(q) relationship can be written as:

$$Y_t = \Theta^q(B)\epsilon_t$$

(in a MA process, it's always stationary. the sum of error terms is also weakly stationary)

Remarks:

- An MA(q) process ia always stationary, regardless of q and the values of the θ's
- An MA(q) process is **"q-correlated"** which means that:

$$\rho(h) = \begin{cases} \neq 0 & if\ h \leq q \\ 0 & if h > q \end{cases}$$

- The opposite is also true: A time series with an ACF that exhibits this pattern can be modeled by an MA(q) model.
- Thus, the ACF plot can be used to identify the order q of an MA process. But it is not helpful in choosing the order p of an AR process.

Partial Autocorrelation

The ACF of lag h measures the correlation between Y_t and Y_{t+h}. This correlation could be due to a direct relationship between Y_t and Y_{t+h}, or it may influenced by observations at the intermediate lags:

$$Y_{t+1}, Y_{t+2}, ...Y_{t+h-1}$$

The PACF of lag h measures the correlation between Y_t and Y_{t+h} ofter accounting the influence of the intermediate lags. We do this by considering:

$$\hat{Y}_t = f(Y_{t+1}, Y_{t+2}, ...Y_{t+h-1})$$
$$\hat{Y}_{t+h} = g(Y_{t+1}, Y_{t+2}, ...Y_{t+h-1})$$

For a time series $\{Y_t\}$ the PACF of lag h is:

$$\alpha(h) = \begin{cases} Corr(Y_t, Y_t) = 1 & if\ h = 0 \\ Corr(Y_t, Y_{t+1} = \rho(1)) & if\ h = 1 \\ Corr(Y_t - \hat{Y}_t, Y_{t+h} - \hat{Y}_{t+h}) & if\ h \geq 2 \end{cases}$$

*

Example: Derive the PACF for an AR(1) process.

$Y_t = \phi Y_{t-1} + \epsilon_t, \{\epsilon_t\} \sim WN(0, \sigma^2)$

for h=0,1 : $\alpha(h) = \begin{cases} 1 & if\ h=0 \\ \rho(1) = \phi & if\ h=1 \end{cases}$

for h=2: $\alpha(2) = Corr(Y_t - f(Y_{t+1}), Y_{t+2} - g(Y_{t+1}))$

$= Corr(Y_t - f(Y_{t+1}), Y_{t+2} - \phi Y_{t+1})$

also known that $\epsilon_t = T_T - \phi Y_{t-1}$

$= Corr(Y_t - f(Y_{t+1}), \epsilon_{t+2})$

$= Corr(Y_t, \epsilon_{t+2}) - Corr(f(Y_{t+1}), \epsilon_{t+2})$

both of these correlations are equal to 0 because you can not depend on error term in the future.

$= 0$

A similar argument shows that $\alpha(h) = 0$ for all $h \geq 2$

$$\alpha(h) = \begin{cases} 1 & if\ h=0 \\ \phi & if\ h=1 \\ 0 & if\ h \geq 2 \end{cases}$$

Remarks:

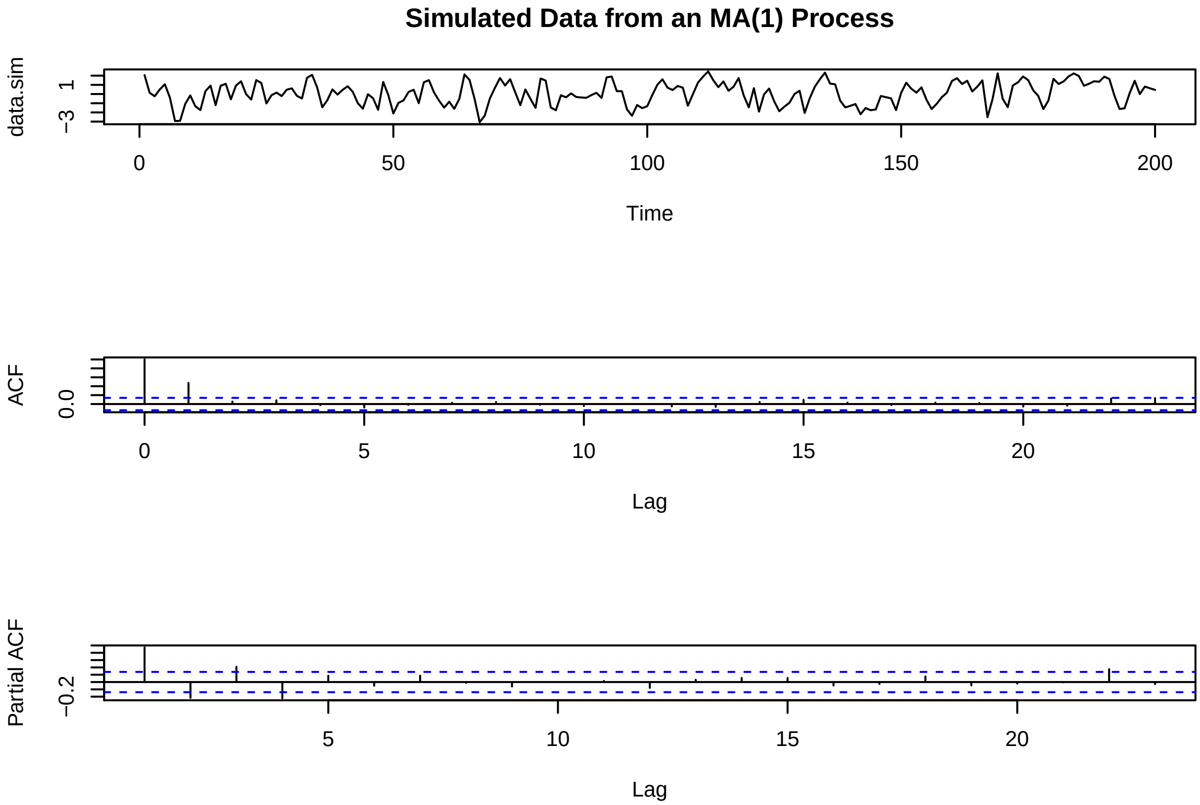
- In general an AR(p) process has a PACF which satisfies:

$$\alpha(h) = \begin{cases} \neq 0 & h \leq p \\ = 0 & h > p \end{cases}$$

- The opposite is also true: if an observed time series has a PACF that exhibits this behaviour then we know it can be modeled by an AR(p) process.

ACF-PACF Examples

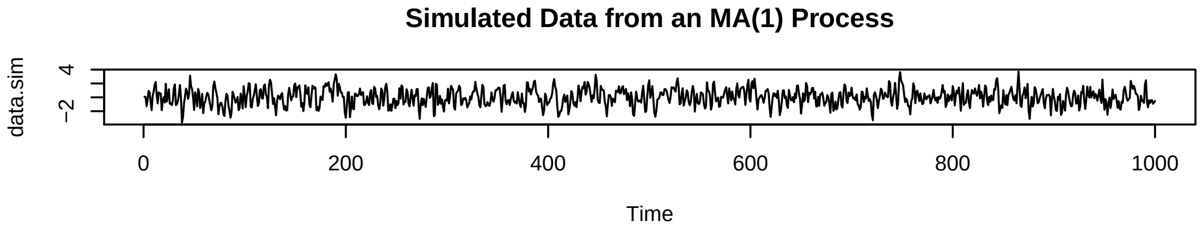
```r
# MA(1) with n = 200
par(mfcol=c(3,1))
data.sim <- arima.sim(n = 200, list(ma = c(0.7)), sd = sqrt(1))
plot(data.sim, main="Simulated Data from an MA(1) Process")
acf(data.sim, main = "")
pacf(data.sim, main = "")
```

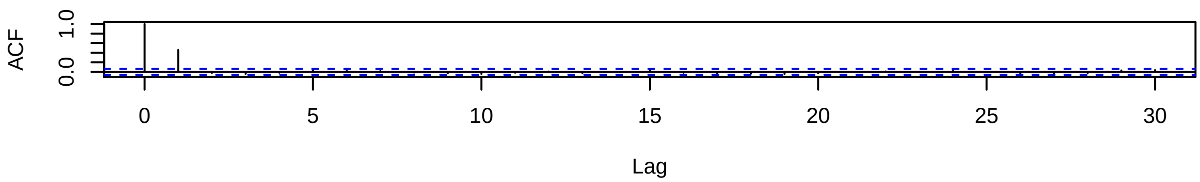


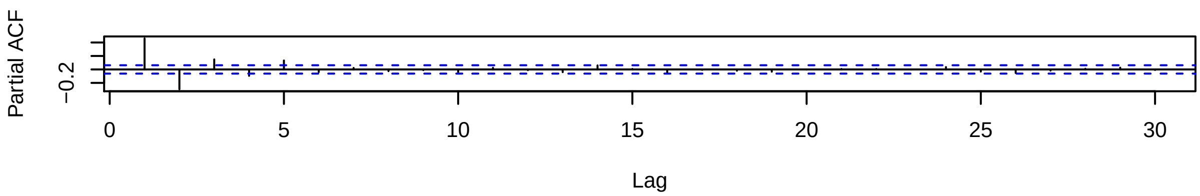

```
# or: acf(data.sim, type="partial")
```

we can observe exponential decay at both sides in PACF. We see 4 significant spikes here.but situation is not obvious here. so we increase the sample size.

```
# MA(1) with n = 1000
par(mfcol=c(3,1))
data.sim <- arima.sim(n = 1000, list(ma = c(0.7)), sd = sqrt(1))
plot(data.sim, main="Simulated Data from an MA(1) Process")
acf(data.sim, main = "")
pacf(data.sim, main = "")
```

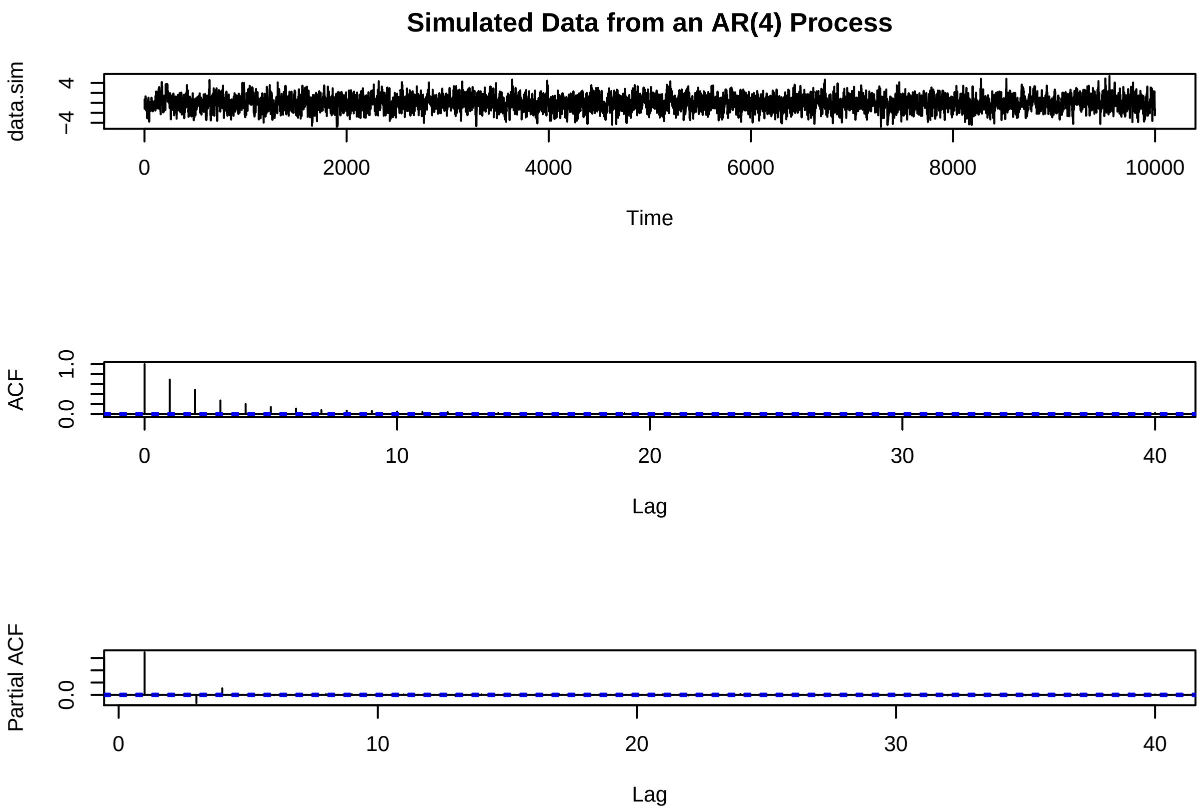






fit all the models and see how they do in test set.

```
# AR(4) with n = 200
par(mfcol=c(3,1))
data.sim <- arima.sim(n = 10000, list(ar = c(0.7, 0.1, -0.2, 0.1)), sd = sqrt(1))
plot(data.sim, main="Simulated Data from an AR(4) Process")
acf(data.sim, main = "")
pacf(data.sim, main = "")
```



exponential decay very obvious. this could be AR(4). For PACF, we observe 4 spikes and then all 0s.