Lecture 5. Autoregressive-Moving Average Models

ARMA

1. Properties

While the Wold form is important for theoretical analysis, it is not practically useful for estimation purposes, since it involves estimating an "infinite" number of parameters $\psi_1, \psi_2, \ldots$. For this reason, the *autoregressive-moving average* (ARMA) models of Box and Jenkins (1976) is popular for modeling a class of stationary processes.

Definition 1.1. A general ARMA(p, q) model is defined by

AR MA

$$x_t = \phi_0 + \sum_{i=1}^{p} \phi_i x_{t-i} + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i},$$

where p is the order of the AR part, q is the order of the MA part, and $\varepsilon_t \sim WN(0, \sigma_\varepsilon^2)$.

- ARMA models approximate the Wold form of a stationary process with a model with a "finite" number of parameters. In specific, the ARMA(p, q) model estimates finite terms $\phi_0, \ldots, \phi_p$ and $\theta_1, \ldots, \theta_q$ for approximating any stationary process. This approximation works well in practice, in the sense that ARMA models with small values of p or q (usually, less than three) are generally sufficient for the analysis of stationary data.

  wold : $x_t \sim MA(\infty)$.

- Using a lag operator, one simplifies the ARMA(p, q) model to

  ARMA(p.q) $\approx$ MA(∞)

  $$\phi(L)x_t = \phi_0 + \theta(L)\varepsilon_t,$$

  where $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \cdots - \phi_p L^p$ and $\theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \cdots + \theta_q L^q$.

Theorem 1.2. *If all of the solutions of the characteristic equation $\phi(z)$ are greater than 1 in modulus, then the ARMA model is stationary. A sufficient condition for invertibility is that all the solutions of the characteristic equation $\theta(z)$ are greater than 1 in modulus.*

Remark 1.3. If x_t follows ARMA(p, q) model and is stationary, one obtains

$$E[x_t] = \mu = \frac{\phi_0}{1 - \phi_1 - \cdots - \phi_p},$$

provided that $\phi_1 + \cdots + \phi_p \neq 1$. So, the alternative representation for the ARMA(p, q) model is

$$(x_t - \mu) = \sum_{i=1}^{p} \phi_i (x_{t-i} - \mu) + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i} \tag{1.1}$$

or

$$\phi(L)(x_t - \mu) = \theta(L)\varepsilon_t.$$

In R, the ARMA(p, q) model is in the form of (1.1) and μ is referred to as the `intercept`.

2. Identifying ARMA Models

In estimating an ARMA(p, q) model, one must decide p and q. A rule of thumb is to use some information criteria.

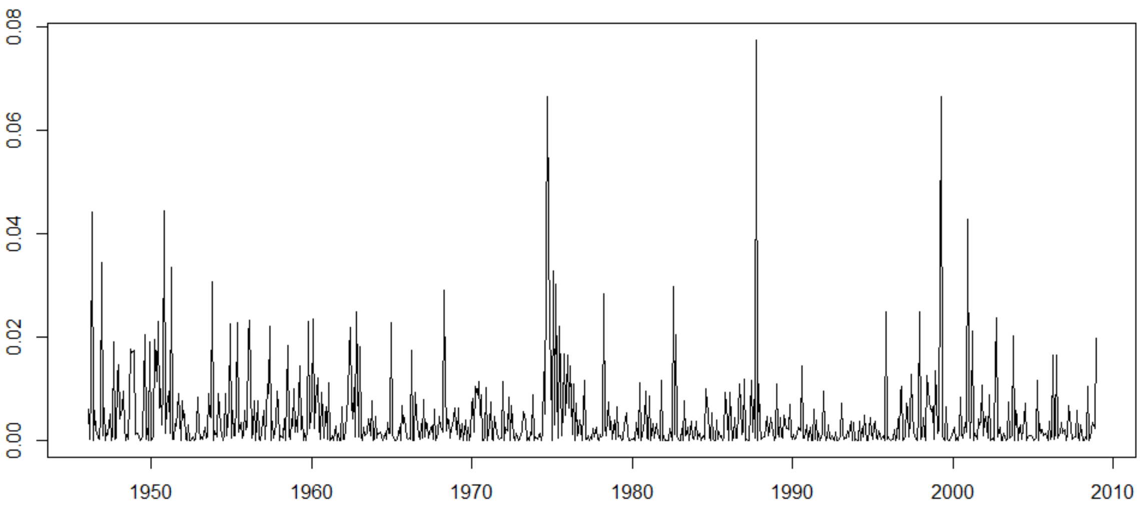
AIC(0,0) (2,2).
(1,0 (2,0) ~ 9개.

- Two common information criteria are the *Akaike Information Criterion* (AIC) and *Schwartz-Bayesian Criterion* (BIC). One computes AIC(k,l) for $k = 0, 1, \dots, p_{max}$ and $l = 0, 1, \dots, q_{max}$, where p_{max} and q_{max} are prespecified positive integers and selects the orders k and l for minimizing the AIC value. The same rule applies to BIC.

Remark 2.1. Once p and q are specified, one can estimate parameters of ARMA models using either the conditional or the exact likelihood method. Indeed, the adequacy of a fitted model is checked via the Ljung-Box statistics of the residuals; i.e., if the model is correctly specified, then $Q(m)$ follows an asymptotic chi-squared distribution with $m - g$ degrees of freedom, where g denotes the number of AR or MA coefficients in the model. Finally, forecasting is done in a recursive manner similar to but a bit more complicated.

Example 2.2. Consider squared monthly returns on 3M stock from February 1946 to December 2008.

```
> mydat <- read.table("data1.txt", header = T)
> head(mydat)
      date       rtn
1 19460228 -0.077922
...
6 19460731  0.076531
> rtn.sq <- ts(mydat$rtn^2, start = c(1946, 2), freq = 12)
> plot(rtn.sq, main ="", xlab = "", ylab = "")
```




```
> Box.test(rtn.sq, lag = 12, type = "Ljung")
```

```
Box-Ljung test
data: rtn.sq
X-squared = 31.2548, df = 12, p-value = 0.001801
```

- For the squared monthly 3M return, the null hypothesis of no autocorrelation is strongly rejected at the 5% level.

```
> computeAIC <- function(p, q){
    reg <- arima(rtn.sq, order = c(p, 0, q), method = "ML")
    reg$aic
    }
> computeAIC(0, 1)
[1] -5232.153
...
> computeAIC(3, 3)
[1] -5237.728
```

Stationarity
→ ARMA
If not, X

- The AIC values computed with $p_{max} = 3$ and $q_{max} = 3$ are tabulated as follows:

	MA			
AR	0	1	2	3
0		-5232.153	-5235.149	-5236.479
1	-5232.697	**-5244.937**	-5243.724	-5241.726
2	-5236.759	-5243.719	-5241.575	-5239.727
3	-5238.443	-5241.725	-5239.722	-5237.728

```
> (reg <- arima(rtn.sq, order = c(1, 0, 1), method = "ML"))
Call:
arima(x = rtn.sq, order = c(1, 0, 1), method = "ML")
Coefficients:
        ar1     ma1  intercept
     0.8871 -0.8155     0.0043
s.e. 0.0571  0.0707     0.0004
sigma^2 estimated as 5.585e-05: log likelihood = 2625.47, aic = -5244.94
> e <- reg$residuals
> mytest <- Box.test(e, lag = 12, type = "Ljung")
> (pvalue <- 1 - pchisq(as.numeric(mytest$statistic), 10))
[1] 0.9759484
```

- The fitted model is

$$(x_t - 0.0043) = 0.8871(x_{t-1} - 0.0043) + \varepsilon_t - 0.8155\varepsilon_{t-1}.$$

For the residual computed from an ARMA(1, 1) model, the null hypothesis of no autocorrelation is not rejected. So, the ARMA(1, 1) model is adequate for squared monthly return on 3M stock.