Nov 6 Notes

Hongdou Li

11/6/2018

Recap

- An AR(p) model is stationary iff $AR(p) = MA(\infty)$
  - This is satisfied **iff** the zeros of the AR generating function lie outside the unit circle in the complex plane.
  - This the **"stationary condition"** for AR models.
- An MA(q) model is only useful if it has an infinite order AR representation. (ie. $MA(q) = AR(\infty)$)
  - This is satisfied **iff** the zeros of the MA generating function lie outside the unit circle in the complex plane.
  - This is the **"Invertibility condition"** for MA models.

ARMA(p,q) Models

$\{Y_t\}$ is an autoregressive moving average of orders p and q if the following conditions are true:

$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + ... + \phi_p Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + ... + \theta_q \epsilon_{t-q}$

where $\{\epsilon_t\} \sim WN(0, \sigma^2)$ and $\phi_1, ...\phi_p, \theta_1 ... \theta_q$ are constants.

This model can be represented in terms of generating functions as follows

$$Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - ... - \phi_p Y_{t-p} = \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + ... + \theta_q \epsilon_{t-q}$$

$$(1 - \phi_1 B - \phi_2 B^2 - ... - \phi_p B^p) Y_t = (1 + \theta_1 B + \theta_2 B^2 + ... + \theta_q B^q) \epsilon_t$$

$$\phi(B) Y_t = \theta(B) \epsilon_t$$

where $\phi(Z) = 1 - \phi_1 Z - \phi_2 Z^2 - ... - \phi_p Z^p$ is the generating function of the AR component.

and $\theta(Z) = 1 + \theta_1 Z + \theta_2 Z^2 + ... + \theta_q Z^q$ is the generating function of the MA component.

Remarks:

- An ARMA(p,q) models is not necessarily stationary, but we'd like it to be so that it could be used to model stationary time series.
- An ARMA(p,q) model is stationary **iff** its AR component is stationary.
  - we check this by determing whether the generating function of the AR component satisfies the stationary conditions.
- An ARMA(p,q) model is not necessarily invertible, but we'd like it to be so that we chould write Y_t exclusively as a function of previous Y's
- An ARMA(p,q) model is invertible **iff** its MA component is invertible.
  - we can check this by determing whether or not the generating function of the MA component satisfies the invertibility conditions.

Other Comments

- AR(p) = ARMA(p,0)
- MA(Q) = ARMA(p,q)
- Model selection in the context of ARMA models require us to choose p and q. We can use ACF and PACF plots to help with this:
  - ACF: q spikes (lags h $\leq$ q) + (mixed) exponential decay
  - PACF: p spikes (lags h $\leq$ q) + (mixed) exponential decay

Example:

$\{Y_t\} \sim ARMA(2,3)$

$Y_t = Y_{t-1} + 0.5Y_{t-2} + \epsilon_t + 0.2\epsilon_{t-1} + 0.7\epsilon_{t-2}$

Determine whether $\{Y_t\}$ is stationary and/or invertible

$Y_t - Y_{t-1} - 0.5Y_{t-2} = \epsilon_t + 0.2\epsilon_{t-1} + 0.7\epsilon_{t-2}$

$(1 - B - 0.5B^2)Y_t = (1 + 0.2B + 0,7B^2)\epsilon_t$

$\phi(B)Y_t = \theta(B)\epsilon_t$

where:

$$\phi(Z) = 1 - Z - 0.5Z^2$$

$$\phi(Z) = 0 \text{ if } Z = \frac{1 \pm \sqrt{(-1)^2 - 4(-0.5)(1)}}{2(-0.5)} = -1 \pm \sqrt{3}$$

$Z_1 = -1 - \sqrt{3} = -2.73$ and $|Z_1| = 2.73$

$Z_2 = -1 + \sqrt{3} = 0.73$ and $|Z_2| = 0.73$

since $|Z_2| < 1$ this root lies inside the unit circle and so the ARMA model is not stationary.

$$\theta(Z) = 1 + 0.2Z + 0,7Z^2$$

$$\theta(Z) = 0 \text{ if } Z = \frac{-0.2 \pm \sqrt{(0.2)^2 - 4(-0.7)(1)}}{2(-0.7)} = \frac{-2 \pm \sqrt{-2.76}}{1.4} = \frac{-2 \pm \sqrt{2.76}i}{1.4}$$

$Z_1 = -0.14 - 1.19i$ and $|Z_1| = \sqrt{(-0.14)^2 + (-1.19)^2} = 1.198$

$Z_2 = -0.14 + 1.19i$ and $|Z_2| = 1.198$

since $|Z_1| = |Z_2| > 1$ this root lies outside the unit circle and so the ARMA model is invertible.

just identify it is an ARMA model, choosing the exact p and q can be very difficult.

```
#data.sim <- arima.sim(n=10000, list(ar=c(0.3,0.65),ma=c(-0.5),sd=sqrt(1)))
#auto.arima(data.sim)
```

Estimating ARMA(p,q) models

Goal: to estimate $\phi_1, \phi_2, ..., \phi_p, \theta_1, \theta_2, ...\theta_q, \sigma$ in the context of an ARMA model:

$\phi(B)Y_t = \theta(B)\epsilon_t$

Those parameters are estimated with observed data $\{y_1, y_2, ...y_n\}$

several methods of estimation exist, but we'll focus on :

(1) Maximum likelihood estimation

(2) Least Squares Estimation