Modeling Strategies (Chapter 4)

Traditional ARMA(n,m) approach

- If either one of m and n is zero and the other is small, the estimated sample correlations are not distorted, one might be able to guess the nonzero order from the cut-off point in the $\hat{\rho}_k$

  e.g., MA(1) model,

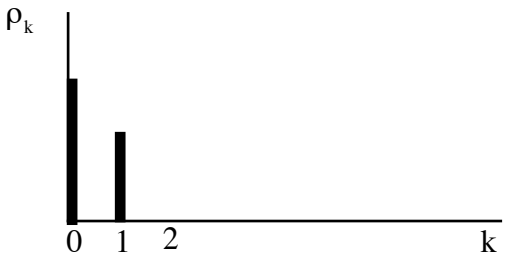
  $X_t = a_t - \theta_1 a_{t-1}$ $\quad G_0=1,\ G_1=-\theta_1,\ G_j=0$ for $j\geq 2$

  $\gamma_0 = \sigma_a^2 - \theta_1 G_1 \sigma_a^2 = (1+\theta_1^2)\,\sigma_a^2$

  $\gamma_1 = -\theta_1 \sigma_a^2$

  $\gamma_\kappa = 0, \quad$ for $k \geq 2$

  $\rho_0=1, \quad \rho_1 = -\dfrac{\theta_1}{1+\theta_1^2}, \quad \rho_\kappa = 0, \quad k \geq 2$



- In general, one may not be able to guess the orders m and n at all or incorrectly. Thus, a trial and error approach is the only choice.
- A rational way of doing this trial and error modeling is to successively fit all the models of orders greater than one, two, and so on.
  e.g., ARMA(5,3) model may require 27 models to be fitted before reaching the adequate one.

System approach by ARMA(n,n-1)

- First, intuitively, as the autoregressive order n increases

  ARMA(n,m): the number of models increase in steps of m.

  ARMA(n,n-1): in steps of one.

- From systems' point of view:
  For a stable system, the dynamic response of the system to a single a will eventually die out to zero.

Let us show why the fancy name "system approach" (new material in Lec. 10).

i) No dynamics:

$G_j \equiv 0,\ j>1, \quad X_t = a_t$

AR(0)

ii) First order dynamics:

very unlikely the influence of a_t at time t+1 would be very large and at the next moment t+2 it would drop to zero.

$G_j = \lambda_1^j \qquad (1-\phi_1 B)X_t = a_t \qquad \lambda_1=\phi_1$

AR(1)

e.g.,

$\lambda_1=\phi_1= 0.1$

$G_0 = 1, \quad G_1= 0.1, \quad G_2= 0.01, \quad G_j\approx 0 \ \text{ for } j \geq 2$

$X_t = a_t - \theta_1 a_{t-1} \qquad \theta_1= -0.1$

MA(1)

$\lambda_1=\phi_1= 0.3$

$$X_t = \frac{a_t}{(1-\phi_1 B)} = (1+0.3B+0.09B^2+0.027B^3+ \ldots)a_t \approx a_t + 0.3a_{t-1}+0.09a_{t-2}+0.027a_{t-3}$$

MA(3)

ii) Second order dynamics:

$G_j = g_1\lambda_1^j + g_2\lambda_2^j$

By allowing λ_1 and λ_2 to be a real or complex conjugate, Green's function can represent different types of dynamics, e.g.,

continuously decreasing memory,

lagged memory,

damped sinusoidal dynamics.

Transition from AR(1) to ARMA(2,1)

AR(1): $\gamma_k = \gamma_0\lambda_1^\kappa$ simple dynamics by single exponential

ARMA(2, 1) $\gamma_\kappa=d_1\lambda_1^\kappa+d_2\lambda_2^\kappa$ weighted sum of two expo.

$\gamma_0 = d_1 + d_2 \qquad d_1 \ \& \ d_2$ restricted $\qquad d_1= \gamma_0 - d_2$

IF the free coefficient d happens to take the value that satisfy

$\gamma_0 = d_1 + d_2$ $\quad d_1$ & d_2 restricted $\quad d_1 = \gamma_0 - d_2$

$\gamma_1 = \phi_1\gamma_0 + \phi_2\gamma_1 \quad (1-\phi_2)\gamma_1 = \phi_1\gamma_0 \quad \gamma_1 = d_1\lambda_1 + d_2\lambda_2$

$(1-\phi_2)[\lambda_1(\gamma_0 - d_2) + \lambda_2 d_2)] = \phi_1\gamma_0$

THEN an AR(2) model will fit the data.

But, in general, this equality is not satisfied. Thus, ARMA(2,1) is required.

$\gamma_1 = \phi_1\gamma_0 + \phi_2\gamma_1 - \theta_1\sigma_a^2$

ARMA(2,1) model also corresponds to a uniformly sampled one-degree-of-freedom vibration system, mass-spring-dashpot system.

e.g., ARMA(2,1) model with $\lambda_1=0.2$ and $\lambda_2=0.4$ $\theta_1=-0.1$

$(1-0.6B+0.08B^2)X_t = (1+0.1B)a_t$

$$X_t = \frac{(1+0.1B)}{(1-0.2B)(1-0.4B)}a_t = (1+0.1B)(1+0.2B+0.04B^2+\ldots)(1+0.4B+0.16B^2+..)a_t$$

$$\approx a_t + 0.7a_{t-1} + 0.34a_{t-2} + 0.148a_{t-3} + 0.061a_{t-4} + 0.025a_{t-5}$$

It becomes a MA(5) model.

iii) Third order dynamics:

$G_j = g_1\lambda_1^j + g_2\lambda_2^j + g_3\lambda_3^j$

In general cases, it will be an ARMA(3,2) model. In special circumstances, it could be any other ARMA(n,m) m<n-1.

iv) nth order dynamics

$G_j = g_1\lambda_1^j + g_2\lambda_2^j + g_3\lambda_3^j + \ldots + g_n\lambda_n^j$

In general cases, it will be an ARMA(n,n-1) model. In special circumstances, it could be any other ARMA(n,m) m<n-1.

Increment in the autoregressive order

- It is theoretically feasible to develop models by increasing n in steps of one. i.e., AR(1), ARMA(2,1), ARMA(3,2),....
  But, it is still not economical in practice when a high order of model is needed.
- ARMA(2n,2n-1) is a better modeling procedure. Besides the empirical experience, there are two reasons for this choice:
  (a) ARMA(2,1) --- one d-0-f dynamic system

ARMA(4,2) --- two d-o-f dynamic system

...

ARMA(2n,2n-1) n d-o-f dynamic system

Thus, increasing the degree-of-freedom by one requires the autoregressive order advance by two.

(b) Configuration of the characteristic roots, λ_i:

Since the autoregressive parameters are always real, the complex roots can occur only in conjugate pairs.

This fact implies that whenever the autoregressive order is odd, one of the roots must be real. If we increase the autoregressive order by one, we force one of the roots to be real.

e.g.,

If an ARMA(4,3) model should be adequate, when we advance from ARMA(2,1) to ARMA(3,2) it might give a poorer model than ARMA(2,1).

Also, increasing the autoregressive order by two is more economical.

2. Checks of Adequacy (new in lecture 10)

Each test can be interpreted as a comparison of two models: one is referred to as the full or complete model; the other is called the reduced model.

ARMA(2n,2n-1) ---> ARMA(2n+2, 2n+1)

H_0: $\phi_{2n+2} = 0, \phi_{2n+1} = 0, \theta_{2n}=0, \theta_{2n+1}=0$

H_1: $\phi_{2n+2} \neq 0, \phi_{2n+1} \neq 0, \theta_{2n}\neq 0, \theta_{2n+1}\neq 0$

In a test of hypothesis, one usually assumes a conservative "null" hypothesis, H_0, and an alternative hypothesis, H_1. If the statistical criterion favors H_1, one concludes that there is evidence against H_0 at a preassigned significance level.

Statistical criteria:

F-test:

$$F = \frac{A_1 - A_0}{s} \div \frac{A_0}{N-r} \quad \sim F(s,\ N-r)$$

A_0 - the (smaller) sum of squares of the unrestricted model;

A_1 - the (larger) sum of squares of the restricted model;

$F(s, N-r)$ -- F-distribution with s and N-r degrees of freedom;

s - number of restricted parameters;

r - number of total unrestricted parameters;

N - number of observations.

e.g., the F-test of ARMA(2n,2n-1) vs. ARMA(2n+2, 2n+1)

$r = (2n+2) + (2n+1) + 1 = 4n + 4$

$s = (4n+4) - [4(n-1) + 4] = 4$

A_0 = RSS of ARMA(2n+2, 2n+1)

A_1 = RSS of ARMA(2n, 2n-1)

$$F = \frac{A_1 - A_0}{4} \div \frac{A_0}{N-4n-4} \quad \sim F(4,\ N-4n-4)$$

if $F > F_{(4,\ N-4n-4)}$ for a given significance level, say, 5%. then the improvement in RSS in going from ARMA(2n,2n-1) to ARMA(2n+2, 2n+1) is significant.

Remarks:

- Theoretically, the F-criterion can be justified only when at least one of the two models is adequate.
- But, the F-criterion may still be used to compare the RSS every time n is increased by one, as long as we get significant F values. In this case, even without the criterion, the reduction in RSS is often so large that one would not decide to stop increasing n.
- To be on the safe side, the decision to stop should be made only after the autocorrelations of the residual a's are also small, $\pm 2/\sqrt{N}$ band or unified autocorrelations within ± 2 band.