Lecture 9: Model Building

Bus 41910, Time Series Analysis, Mr. R. Tsay

An effective procedure for building empirical time series models is the Box-Jenkins approach, which consists of three stages: model specification, estimation and diagnostics checking. These three stages are used iteratively until an appropriate model is found. The estimation is accomplished by using mainly the maximum likelihood method. For model checking, there are various methods available in the literature, and we shall discuss some of those methods later. For now, we shall focus on model specification.

Model specification (or identification) is intended to specify, from the data, certain tentative models which are worth a careful investigation. For simplicity, we focus on the class of ARIMA models. However, the three-stage modeling procedure applies equally well to other models. For ARIMA models, there are two main approaches to model specification. The first approach is called the "correlation" approach in which the tentative models are selected via the examination of certain (sample) correlation functions. This approach does not require "full" estimation" of any model. However, it is judgemental in the sense that a data analyst must make a decision regarding which models to entertain. The second approach is called the information criterion approach in which an objective function is defined and the model selection is done automatically by evaluating the objective function of possible models. Usually, the model which achieves the minimum of the criterion function is treated as the "most appropriate" model for the data. The evaluation of the criterion function for a given model, however, requires formal estimation of the model.

Suppose that the observed realization is $\{Z_1, Z_2, \cdots, Z_n\}$. In some cases, certain transformation of Z_t is needed before model building, e.g. variance stablization. Thus, one should always plot the data before considering model specification. In what follows, we shall briefly discuss the two model-specification approaches.

A. Correlation approach: The basic tools used in this approach of model specification include (a) sample autocorrelation function (ACF), (b) sample partial autocorrelation function (PACF), (c) extended autocorrelation function (EACF) and (d) the method of smallest canonical correlation (SCAN). The function of these tools can be summarized as

Function	Model	Feature
ACF	MA(q)	Cutting-off at lag q
PACF	AR(p)	Cutting-off at lag p
EACF	ARMA(p, q)	A triangle with vertex (p, q)
SCAN	ARMA(p, q)	A rectangle with vertex (p, q)

Illustration:

a. <u>ACF</u>: The lag-ℓ sample ACF of Z_t is defined by

$$\hat{\rho}_\ell = \frac{\sum_{t=\ell+1}^{n}(Z_t - \bar{Z})(Z_{t-\ell} - \bar{Z})}{\sum_{t=1}^{n}(Z_t - \bar{Z})^2}$$

where $\bar{Z} = \frac{1}{n}\sum_{t=1}^{n} Z_t$ is the sample mean. In the literature, you may see some minor deviation from this definition. However, the above one is close to being a standard. Two main features of sample ACF are particularly useful in model specification. First of all, for a stationary ARMA model,

$$\hat{\rho}_\ell \rightarrow_p \rho_\ell, \quad \text{as} \quad n \rightarrow \infty$$

where $\rightarrow_p$ denotes convergence in probability. Also, $\hat{\rho}_\ell$ is asymptotically normal with mean ρ_ℓ and variance being function of the ACF ρ_i's. (See Box and Jenkins (1976) and the references therein. Or page 21 of Wei (1990)). Recall that for an MA(q) process, we have

$$\rho_\ell \begin{cases} \neq 0 & \text{for } \ell = q \\ = 0 & \text{for } \ell > q. \end{cases}$$

Therefore, for moderate and large samples, the sample ACF of an MA(q) process would show this cutting-off property. In other words, if

$$\hat{\rho}_q :\neq 0, \quad \text{but} \quad \hat{\rho}_\ell := 0 \quad \text{for} \quad \ell > q,$$

then the process is likely to follow an MA(q) model. Here $:=$ and $:\neq$ denote, respectively, statistically equal to and different from. To judge the significance of sample ACF, we use its asymptotic variance under certain null-hypothesis. It can be shown that for an MA(q) process, the asymptotic variance of $\hat{\rho}_\ell$ for $\ell > q$ is

$$V[\hat{\rho}_\ell] = \frac{1 + 2(\rho_1^2 + \cdots + \rho_q^2)}{n}.$$

This is referred to as the Bartlett's formula in the literature. See Chapter 6, page 177, of Box and Jenkins (1976). In practice, the ρ_i's are estimated by $\hat{\rho}_i$'s. In particular, if Z_t is a white noise process, than $V[\hat{\rho}_\ell] = 1/n$ for all $\ell > 0$. See the computer output of ACF.
The second important feature of sample ACF is that for any ARIMA(p, d, q) model with $d > 0$,

$$\hat{\rho}_\ell \rightarrow_p 1 \quad \text{as} \quad n \rightarrow \infty.$$

This says that the sample ACF is persistent for any ARIMA(p, d, q) model. In practice, persistent sample ACF is often regarded as an indication of non-stationarity and differencing is used to render the series stationary. See computer output on differencing.

b. <u>PACF</u>: Recall that ACF of an ARMA(p, q) model satisfies $\phi(B)\rho_\ell = 0$ for $\ell > q$. In particular, for AR models, the ACF satisfies the difference equation $\phi(B)X = 0$, implying that the ACF has infinite non-zero lags and tends to be damped sine (co-sine) function or exponentials. Thus, sample ACF is not particularly useful in specifying pure AR models.

On the other hand, recall that the Yule-Walker equation of an AR(p) process can be used to obtain the AR coefficients from the ACF. Obviously, for an AR(p) model, all the AR-coefficients of order higher than p are zero. Consequently, by examining the estimates of AR coefficients, one can identify the order of an AR process. The p-th order Yule-walker equation is

$$\begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{p-2} & \rho_{p-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{p-3} & \rho_{p-2} \\ \vdots & & & & & \vdots \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \cdots & \rho_1 & 1 \end{bmatrix} \begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{bmatrix}.$$

By the Cramer rule, we have

$$\phi_p = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{2-p} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{3-p} & \rho_2 \\ \vdots & & & & & \vdots \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \cdots & \rho_1 & \rho_p \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{p-2} & \rho_{p-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{p-3} & \rho_{p-2} \\ \vdots & & & & & \vdots \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \cdots & \rho_1 & 1 \end{vmatrix}}. \tag{1}$$

Let $\hat{\phi}_{p,p}$ be the estimate of ϕ_p obtained via equation (1) with ρ_ℓ replaced by its sample counterpart $\hat{\rho}_\ell$. The function

$$\hat{\phi}_{1,1}, \quad \hat{\phi}_{2,2}, \quad \cdots, \quad \hat{\phi}_{\ell,\ell}, \quad \cdots$$

is called the sample PACF of Z_t. Based on the previous discussion, for an AR(p) process, we have

$$\hat{\phi}_{p,p} :\neq 0, \quad \text{but} \quad \hat{\phi}_{\ell,\ell} := 0 \quad \text{for} \quad \ell > p.$$

This is the cutting-off property of sample PACF by which the order of an AR process can be specified.

Alternatively, the sample PACF $\hat{\phi}_{\ell,\ell}$ can be defined as the least squares estimates of the following consecutive autoregressions:

$$\begin{aligned} Z_t &= \phi_{1,0} + \phi_{1,1} Z_{t-1} + e_{1t} \\ Z_t &= \phi_{2,0} + \phi_{2,1} Z_{t-1} + \phi_{2,2} Z_{t-2} + e_{2t} \\ Z_t &= \phi_{3,0} + \phi_{3,1} Z_{t-1} + \phi_{3,2} Z_{t-2} + \phi_{3,3} Z_{t-3} + e_{3t} \\ \vdots &= \vdots \end{aligned}$$

This later explanation is more intuitive. It also works better when the process Z_t is an ARIMA(p, d, q) process. The first definition of sample PACF via sample ACF is not well-defined in the case of ARIMA processes. The two definitions, of course, are the same in theory when the series Z_t is stationary.

In practice, it can be shown that the for an AR(p) process, the asymptotic variance of the sample PACF $\hat{\phi}_{\ell,\ell}$ is $\frac{1}{n}$ for $\ell > p$. See computer output.

c. EACF. The model specification of mixed ARMA models is much more complicated than that of pure AR or MA models. We shall consider two methods. The first method to identify the order of a mixed model is the extended autocorrelation function (EACF) of Tsay and Tiao (1984, JASA). (You are recommended to read the article for details.) The EACF, in fact, applies to ARIMA as well as ARMA models. However, it treats an ARIMA(p, d, q) model as an ARMA($p + d, q$) model.

The basic idea of EACF is based on the "generalized" Yule-Walker equation. Conceptually, it involves two steps. In the first step, we attempt to obtain consistent estimates of AR coefficients. Given such estimates, we can transform the ARMA series into a pure MA process. The second step then uses the sample ACF of the transformed MA process to identify the MA order q.

The best way to introduce EACF is to consider some simple examples.

Example 1: Suppose that Z_t is an ARMA(1,1) model

$$Z_t - \phi Z_{t-1} = a_t - \theta a_{t-1}, \quad |\phi| < 1, \quad |\theta| < 1.$$

For this model, the ACF is

$$\rho_\ell = \begin{cases} \frac{(1-\phi\theta)(\phi-\theta)}{1+\theta^2-2\phi\theta} & \text{for } \ell = 1 \\ \phi\rho_{\ell-1} & \text{for } \ell > 1. \end{cases}$$

For $p = 1$, the usual Yule-Walker equation is

$$\rho_1 = \phi\rho_0,$$

and the j-th generalized Yule-Walker equation is

$$\rho_{j+1} = \phi\rho_j.$$

Denote the solution of the Yule-Walker equation by $\phi_{1,1} = \phi_{1,1}^{(0)}$ and that of the j-th generalized Yule-Walker equation by $\phi_{1,1}^{(j)}$. Then, we have

$$\phi_{1,1}^{(j)} = \begin{cases} \rho_1 \neq \phi & \text{for } j = 0 \\ \phi & \text{for } j > 0 \end{cases}$$

Thus, the solution of the usual Yule-Walker equation is not consistent with the AR coefficient ϕ. However, **ALL** of the solutions of the j-th generalized Yule-Walker equations are consistent with the AR coefficient. In sample, these results say that the estimates of $\phi_{1,1}^{(j)}$ obtained by replacing the ACF by sample ACF have the property:

$$\hat{\phi}_{1,1}^{(j)} \to_p \begin{cases} \rho_1 & \text{for } j = 0 \\ \phi & \text{for } j > 0. \end{cases}$$

Now define the transformed series $W_{1,t}^{(j)}$ by

$$W_{1,t}^{(j)} = Z_t - \hat{\phi}_{1,1}^{(j)} Z_{t-1} \quad \text{for} \quad j > 0.$$

The above discussion shows that $W_{1,t}^{(j)}$ for $j > 0$ is asymptotically a pure MA(1) process. Consequently, by considering the ACF of the $W_{1,t}^{(j)}$ series, we can identify that the MA order is 1.

Example 2: Suppose now that Z_t is a stationary and invertible ARMA(1,2) process

$$Z_t - \phi Z_{t-1} = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2}.$$

The ACF of Z_t satisfies

$$\rho_\ell \begin{cases} \neq \phi\rho_1 & \text{for } \ell = 2 \\ = \phi\rho_{\ell-1} & \text{for } \ell > 2 \end{cases}$$

Using this result and considering the solution of the j-th generalized Yule-Walker equation of order 1

$$\rho_{j+1} = \phi\rho_j,$$

we see that

$$\phi_{1,1}^{(j)} \begin{cases} \neq \phi & \text{for } j \leq 2 \\ = \phi & \text{for } j > 2 \end{cases}$$

Therefore, the j-th transformed series

$$W_{1,t}^{(j)} = Z_t - \phi_{1,1}^{(j)} Z_{t-1}$$

is an MA(2) series provided that $j > 2$.

Compared with the result of Example 1, we see that the difference between ARMA(1,1) and ARMA(1,2) is that we NEED to consider one step further in the generalized Yule-Walker equation. In either case, however, the ACF of the transformed series can suggest the MA order q once a consistent AR coefficient is used.

In general, the above two simple examples show that for an ARMA(1,q) model, the j-th generalized Yule-Walker equation provides consistent AR estimate if $j > q$. Thus, the j-th transform series $W_{1,t}^{(j)} = Z_t - \phi_{1,1}^{(j)} Z_{t-1}$ is an MA(q) series for $j > q$. In practice, it would be cumbersome to consider ACF of all the transformed series $W_{1,t}^{(j)}$ for $j = 1, 2, \cdots$. We are thus led to consider a summary of the ACF. The EACF is a device which is designed to summarize the pattern of ACF of $W_{1,t}^{(j)}$ for all j.

First-order extended ACF: The first-order extended ACF is defined as

$$\rho_{1,j} = \rho_\ell \quad \text{of} \quad W_{1,t}^{(j)}$$

where

$$W_{1,1}^{(j)} = Z_t - \phi_{1,1}^{(j)} Z_{t-1}, \quad \text{with} \quad \phi_{1,1}^{(j)} = \frac{\rho_{j+1}}{\rho_j}, \quad j \geq 0.$$

It is easy to check that for an ARMA(1,q) process, we have

$$\rho_{1,j} \begin{cases} \neq 0 & \text{for } j \leq q \\ = 0 & \text{for } j > q. \end{cases}$$

In summary, the first-order extended autocorrelation function is designed to identify the order of ARMA(1,q) model. It function in an exact manner as that of ACF to an MA model.

Similarly, we can define a 2nd-order EACF to identify the order of an ARMA(2,q) model,

$$Z_t - \phi_1 Z_{t-1} - \phi_2 Z_{t-2} = c + a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}.$$

More specifically, the j-th generalized Yule-Walker equation of order 2 is defined by

$$\begin{bmatrix} \rho_{j+1} \\ \rho_{j+2} \end{bmatrix} = \begin{bmatrix} \rho_j & \rho_{j-1} \\ \rho_{j+1} & \rho_j \end{bmatrix} \begin{bmatrix} \phi_{2,1}^{(j)} \\ \phi_{2,2}^{(j)} \end{bmatrix}.$$

Obviously, the solution of this equation satisfies

$$\phi_{2,i}^{(j)} = \phi_i \quad i = 1, 2; \quad \text{for} \quad j > q.$$

Define the 2nd-order EACF by

$$\rho_{2,j} = \rho_j \quad \text{of the transformed series} \quad W_{2,t}^{(j)}$$

where

$$W_{2,t}^{(j)} = Z_t - \phi_{2,1}^{(j)} Z_{t-1} - \phi_{2,2}^{(j)} Z_{t-2}.$$

It is clear from the above discussion that

$$\rho_{2,j} \begin{cases} \neq 0 & \text{for } j = q \\ = 0 & \text{for } j > q. \end{cases}$$

Here, of course, Z_t is an ARMA(2,q) process.

You should be able to generalize the EACF to the general ARMA(p, q) case. (Exercise!)

Model Specification via EACF. To make use of the EACF for model specification, we consider the two-way table:

AR	MA (or j)					
m	0	1	2	3	4	$\cdots$
0	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	$\cdots$
1	$\rho_{1,1}$	$\rho_{1,2}$	$\rho_{1,3}$	$\rho_{1,4}$	$\rho_{1,5}$	$\cdots$
2	$\rho_{2,1}$	$\rho_{2,2}$	$\rho_{2,3}$	$\rho_{2,4}$	$\rho_{2,5}$	$\cdots$
3	$\rho_{3,1}$	$\rho_{3,2}$	$\rho_{3,3}$	$\rho_{3,4}$	$\rho_{3,5}$	$\cdots$
$\vdots$	$\vdots$			$\vdots$		

The EACF Table

In practice, the EACF in the above table is replaced by its sample counterpart. To identify the order of an ARMA model, we need to understand the behavior of the EACF table for a given model. Before giving the theory, I shall illustrate the function of the table. Suppose that Z_t is an ARMA(1,1) model, then the corresponding EACF table is

AR	MA (or j)						
m	0	1	2	3	4	5	$\cdots$
0	X	X	X	X	X	X	$\cdots$
1	X	O	O	O	O	O	$\cdots$
2	*	X	O	O	O	O	$\cdots$
3	*	*	X	O	O	O	$\cdots$
4	*	*	*	X	O	O	$\cdots$

The EACF Table
where "X" and "O" denotes non-zero and zero quantities, respectively, "*" represents a quantity which can assume any value between -1 and 1.

From the table, we see that there exists a triangle of "O" with vertex at $(1,1)$, which is the order of Z_t. In practice, the non-zero and zero terms are determined by the sample EACF and its estimated standard error via the Bartlett's formula for MA models. Of course, we cannot expect to see an exact triangle as that of the above table. However, one can often make a decision based on the pattern of the EACF table.

To understand the triangular pattern, it is best to consider a simple example such as ARMA(1,1) model of the above table. In particular, we shall discuss the reason why $\rho_{2,2}$ is different from zero for an ARMA(1,1) model. By definition, $\rho_{2,2}$ is the lag-2 ACF of the transformed series

$$W_{2,t}^{(2)} = Z_t - \phi_{2,1}^{(2)} Z_{t-1} - \phi_{2,2}^{(2)} Z_{t-2}$$

where $\phi_{2,1}^{(2)}$ and $\phi_{2,2}^{(2)}$ are the solution of the 2nd generalized Yule-Walker equation of order 2, namely

$$\left[\begin{array}{c} \rho_3 \\ \rho_4 \end{array} \right] = \left[\begin{array}{cc} \rho_2 & \rho_1 \\ \rho_3 & \rho_2 \end{array} \right] \left[\begin{array}{c} \phi_{2,1}^{(2)} \\ \phi_{2,2}^{(2)} \end{array} \right].$$

However, for an ARMA(1,1) model, $\rho_j = \phi\rho_{j-1}$ for $j > 1$ so that the above Yule-Walker equation is "singular" in theory. In practice, the equation is not exactly singular, but is ill-conditioned. Therefore, the solution $\hat{\phi}_{2,1}^{(2)}$ and $\hat{\phi}_{2,2}^{(2)}$ can assume any real numbers. Consequently, the chance that $\phi_{2,i}^{(2)} = 0$ is essential zero. More importantly, this implies that the transformed series $W_{2,t}^{(2)}$ is not an MA(1) series. Therefore, $\rho_{2,2} \neq 0$. Intuitively, one can interpret this result as an over-fitting phenomenon. Since the true model is ARMA(1,1) and we are fitting an AR(2) polynomial in the construction of $W_{2,t}^{(2)}$, the non-zero $\rho_{2,2}$ is in effect a result of overfitting of the second AR coefficient.

Using exactly the same reasoning, one can deduce the triangular pattern of the EACF table. Thus, it can be said that the triangular pattern of EACF is related to the overfitting of AR polynomials in constructing the transformed series $W_{m,t}^{(j)}$.
<u>Illustration</u>:

D. <u>SCAN</u>. Next we consider the SCAN method which is closely related to the EACF approach as both methods rely on the generalized moment equations of a time series. However, the SCAN approach utilizes the generalized moment equations in a different way so that it does not encounter the overfitting problem of EACF. In practice, my experience indicates that EACF tends to specify mixed ARMA models whereas SCAN prefers AR type of models.
Although the SCAN approach applies to the non-stationary ARIMA models, we shall only consider the stationary case in this introduction. The moment equations of an ARMA(p, q) process is

$$\rho_\ell - \phi_1\rho_{\ell-1} - \cdots - \phi_p\rho_{\ell-p} = f(\boldsymbol{\theta}, \boldsymbol{\phi}, \sigma_a^2), \quad \ell \geq 0,$$

where $f(.)$ is a function of its arguments. In particular, for $\ell > q$, we have

$$\rho_\ell - \phi_1\rho_{\ell-1} - \cdots - \phi_p\rho_{\ell-p} = 0. \tag{2}$$

Obviously, Yule-Walker equations and their generalizations are ways to exploit the above moment equation. An alternative approach to make use of the equation (2) is to consider the signularity of the matrices $\boldsymbol{A}(m, j)$ for $m \geq 0$ and $j \geq 0$, where

$$\boldsymbol{A}(m,j) = \left|\begin{array}{ccccc} \rho_{j+1} & \rho_j & \cdots & \rho_{j+2-m} & \rho_{j+1-m} \\ \rho_{j+2} & \rho_{j+1} & \cdots & \rho_{j+3-m} & \rho_{j+2-m} \\ \vdots & & & & \vdots \\ \rho_{j+1+m} & \rho_{j+m} & \cdots & \rho_{j+2} & \rho_{j+1} \end{array}\right|_{(m+1)\times(m+1)}.$$

For example, suppose that Z_t is ARMA(1,1), then

$$\rho_\ell - \phi_1\rho_{\ell-1} = 0 \quad \text{for} \quad \ell > 1.$$

Consequently, by arranging the $\boldsymbol{A}(m, j)$ in a two-way table

AR	MA					
	j					
m	0	1	2	3	4	$\cdots$
0	$\boldsymbol{A}(0,0)$	$\boldsymbol{A}(0,1)$	$\boldsymbol{A}(0,2)$	$\boldsymbol{A}(0,3)$	$\boldsymbol{A}(0,4)$	$\cdots$
1	$\boldsymbol{A}(1,0)$	$\boldsymbol{A}(1,1)$	$\boldsymbol{A}(1,2)$	$\boldsymbol{A}(1,3)$	$\boldsymbol{A}(1,4)$	$\cdots$
2	$\boldsymbol{A}(2,0)$	$\boldsymbol{A}(2,1)$	$\boldsymbol{A}(2,2)$	$\boldsymbol{A}(2,3)$	$\boldsymbol{A}(2,4)$	$\cdots$
$\vdots$						

we obtain the pattern

	j					
m	0	1	2	3	4	$\cdots$
0	N	N	N	N	N	$\cdots$
1	N	S	S	S	S	$\cdots$
2	N	S	S	S	S	$\cdots$
3	N	S	S	S	S	$\cdots$
$\vdots$	$\vdots$					

where N and S denote, respectively, singular and non-singular matrix.

From the table, we see that the order (1,1) corresponds exactly to the vertex of a rectangle of singular matrices.

Mathematically, there are many ways to show singularity of a matrix. For instance, one can use determinant or the smallest eigenvalue. An important consideration here is, of course, the statistical properties of the test statistic used to check singularity of a sample matrix. The SCAN approach makes use of the idea of "canonical correlation analysis", which is a standard technique in multivariate analysis. See, for instance, Anderson (1984). It turns out that there are other advantages in using canonical correlation analysis. For instance, the approach also applies to multivariate time series analysis, see Tiao and Tsay (1989, JRSSB).

For a time series Z_t, the matrix $\boldsymbol{A}(m,j)$ is the covariance matrix between the vectors $\boldsymbol{Y}_{m,t} = (Z_t, Z_{t-1}, \cdots, Z_{t-m})'$ and $\boldsymbol{Y}_{m,t-j-1} = (Z_{t-j-1}, Z_{t-j-2}, \cdots, Z_{t-j-1-m})'$. The singularity of $\boldsymbol{A}(m,j)$ means that a linear combination of $\boldsymbol{Y}_{m,t}$ is uncorrelated with the vector $\boldsymbol{Y}_{m,t-j-1}$. Thinking in this way, it is then easy to understand the SCAN approach.
Let F_t denote the information available up to and including Z_t. In other words, F_t is the σ-field generated by $\{Z_t, Z_{t-1}, Z_{t-2}, \cdots\}$. Then, the equation of an ARMA(p,q) model

$$Z_t - \phi_1 Z_{t-1} - \cdots - \phi_p Z_{t-p} = a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}$$

says, essentially, that the linear combination

$$Z_t - \phi_1 Z_{t-1} - \cdots - \phi_p Z_{t-p} \stackrel{def}{=} (1, -\phi_1, -\phi_2, \cdots, -\phi_p)\boldsymbol{Y}_{p,t}$$

is uncorrelated with F_{t-j-1} for all $j \geq q$. Therefore, for an ARMA(p,q) series, a linear combination of $\boldsymbol{Y}_{p,t}$ is uncorrelated with $\boldsymbol{Y}_{p,t-j-1}$ for all $j \geq q$.
In practice, to test that a linear combination of $\boldsymbol{Y}_{m,t}$ is uncorrelated with $\boldsymbol{Y}_{m,t-j-1}$, the SCAN approach uses the test statistic

$$c(m,j) = -(n-m-j)\log(1 - \frac{\lambda^2(m,j)}{d(m,j)})$$

where n is the sample size, $\lambda^2(m,j)$ is the square of the smallest canonical correlation between $\boldsymbol{Y}_{m.t}$ and $\boldsymbol{Y}_{m,t-j-1}$ and $d(m,j)$ is defined by

$$d(m,0) = 1, \quad d(m,j) = 1 + 2\sum_{k=1}^{j} \hat{\rho}_k^2(W), \quad j > 0$$

where W_t is a transformed series of Z_t based on the eigenvector of $\boldsymbol{A}(m, j)$ corresponding to $\lambda^2(m, j)$. This statistic $c(m, j)$ follows asymptotically a chi-square distribution with 1 degree of freedom for (a) $m = p$ and $j \geq q$ or (b) $m \geq p$ and $j = q$. For further details, see Tsay and Tiao (1985, Biometrika).

Illustration:

Remark: I assume that most of you have the idea of canonical correlation analysis. If you don't, please consult any textbook of multivariate analysis. For example, Anderson (1984) and Mardia, Kent, and Bibby (1979). Roughly speaking, consider two vector variables $\boldsymbol{X}$ and $\boldsymbol{Y}$. Canonical correlation analysis is a technique intended to answer the following questions:

- Q1: Can you find a linear combination of $\boldsymbol{X}$, say $x_1 = \boldsymbol{\alpha}_1'\boldsymbol{X}$, and a linear combination of $\boldsymbol{Y}$, say $y_1 = \boldsymbol{\beta}_1'\boldsymbol{Y}$, such that the correlation between x_1 and y_1 is the maximum among all possible linear combinations of $\boldsymbol{X}$ and all possible linear combinations of $\boldsymbol{Y}$?

- Q2: Can you find a linear combination of $\boldsymbol{X}$, say $x_1 = \boldsymbol{\alpha}_2'\boldsymbol{X}$, which is orthogonal to x_1, and a linear combination of $\boldsymbol{Y}$, say $y_2 = \boldsymbol{\beta}_2'\boldsymbol{Y}$, which is orthogonal to y_1, such that the correlation between x_2 and y_2 is the maximum among all linear combinations of $\boldsymbol{X}$ and all linear combinations of $\boldsymbol{Y}$ that satisfy the orthogonality condition?

Obviously, one can continue the question until the dimenion of $\boldsymbol{X}$ or that of $\boldsymbol{Y}$ is reached. The solutions of the above questions for $\boldsymbol{X}$ trun out to be the eigenvalues and their corresponding eigenvectors of the matrix:

$$[V(X)]^{-1}Cov(X, Y)[V(Y)]^{-1}Cov(Y, X)$$

with the maximum eigenvalue giving rise to the maximum correlation. By interchanging $\boldsymbol{X}$ and $\boldsymbol{Y}$, we obtain the linear combinations of $\boldsymbol{Y}$.