Eric Slud, STAT 730 May 8, 2017

State Space Represenations for ARMA Processes

The book (Shumway and Stoffer) discusses this topic in Sections 6.1 and 6.2. The purpose of this handout is to emphasize that there are at least three different state space representations of stationary and causal $ARMA(p,q)$ processes in common use, and to distinguish their form and features with reference to the **R** Kalman filtering and forecasting functions. In each case, we want to represent the process Y_t satisfying

$$\phi(B)\,Y_t \;=\; \theta(B)\,W_t \quad , \qquad W_t \sim WN(0,\sigma^2) \tag{1}$$

with polynomials $\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$, $\;\theta(z) = 1 + \theta_1 z + \cdots \theta_q z^q$.

Method 1. The simplest one to explain and justify is based on a state-vector $\mathbf{X}_t = (X_t, X_{t-1}, \ldots, X_{t-r+1})'$ of dimension $\;r = \max(p, q+1)$. Denote

$$\underline{\phi} = (\phi_1, \ldots, \phi_r)'\,,\; \underline{\theta} = (\theta_0, \theta_1, \ldots, \theta_r)'\,, \quad \phi_j = 0 \;\text{ for }\; j > p,\;\; \theta_0 = 1,\;\; \theta_j = 0 \;\text{ for }\; j > q$$

The *observation equation* connecting $\mathbf{X}_t$ and Y_t by

$$Y_t \;=\; \underline{\theta}'\,\mathbf{X}_t = X_t + \theta_1 X_{t-1} + \cdots \theta_q X_{t-q}$$

and X_t is defined as the causal $AR(p)$ process associated with $\underline{\phi}$ and W_t, i.e., $\phi(B)X_t = W_t$ or

$$\mathbf{X}_{t+1} \;=\; \mathbf{T}\,\mathbf{X}_t \,+\, \binom{1}{\mathbf{0}} W_t \quad , \quad \text{and } \; T_{1j} = \phi_j,\;\; T_{i,j} = \delta_{i-1,j} \;\text{ for }\; i > 1,\;\; \text{all }\; j \tag{2}$$

The justification of this state equation is that $Y_t \;=\; \theta(B)\,(\phi(B))^{-1}W_t$ while $X_t = (\phi(B))^{-1}W_t$.

Method 2. A second and different representation, which seems to be the one that the **R** function `KalmanForecast` relies on (with reference to a 1980 *Applied Statistics* paper of Gardner, Harvey and Phillips), is based on the same dimension r and transition matrix $\mathbf{T}$, but now with state vector

$$\underline{\alpha}_t \;=\; (\alpha_{t,1}, \ldots, \alpha_{t,r})' \quad , \qquad \alpha_{t,i} \;=\; \sum_{j=i}^{r} \Big[\phi_j\, Y_{t+i-j-1} \,+\, \theta_{j-1}\, W_{t+i-j}\Big] \tag{3}$$

In this case, with observation equation $\;Y_t \;=\; \alpha_{t,1},\;$ it is easy to see for all $i \le r$ that the state-equation holds in the form

$$\alpha_{t,i} \;=\; \phi_i\, \alpha_{t-1,1} \,+\, \theta_{i-1} W_t \,+\, \alpha_{t-1,i+1}$$

where $\alpha_{s,r+1} \equiv 0$ by convention. This last equation can be re-written in matrix form, in terms of the transpose of the same matrix T appearing in *Method 1*, as follows

$$\underline{\alpha}_{t+1} = T' \, \underline{\alpha}_t \, + \, W_t \, (\theta_0, \theta_1, \ldots, \theta_{r-1})' \tag{4}$$

Method 3. Whenever $p \le q$, the dimension r of the previous two state-space representations is $q+1$. However, as presented in the 1991 book of Brockwell and Davis, Example 12.1.6 in Chapter 12, there is a valid general representation of dimension $m = \max(p, q)$. The justification is given there, and is less intuitive. We do not reproduce the definition here.

Initialization and Application to Forecasting

Methods 1 and 2 differ in one important respect from the point of view of forecasting. If observations Y_t, Y_{t-1}, $\ldots Y_1$ are specified, and all Y_s for $s \le 0$ are set to 0 then the conditional expectations of all future Y_s observations ($s > t$) are well-specified given the past through the recursion in *Method 1*, but the same is not true for *Method 2* because the W_s observations are not fixed for $s \le t$. This is an issue of initialization. In the setting of *Method 1* when the *ARMA* process is *invertible*, $Y_s \equiv 0$ for all $s \le 0$ if and only if $X_s \equiv 0$ for all $s \le 0$. In that case, we easily solve the equations

$$X_1 = Y_1, \; X_2 + \theta_1 X_1 = Y_2, \; \ldots \; X_r + \theta_1 X_{r-1} + \cdots + \theta_q X_{r-q} = Y_r$$

recursively, and then for all $s > r$, the observation equation determines further X_s values from further Y_s values (when none are missing).

We indicate in the *.RLog files on the course web-page, first in the $AR(p)$ case, how to do forecasting and standardized residuals from ARMA model fits using the `KalmanForecast` function in **R**.