Extending AR Models for Complex Time Series Data

After finish learning the materials from Module 1, you may wonder why we start with AR models. As we mentioned at the beginning, AR models are central to stationary time series data analysis and, as components of larger models or in suitably modified and generalized forms, underlie nonstationary time-varying models. This reading material will give examples about extending AR models for complex time series data. Hopefully this notes will help you build up confidence to this relatively simple model.

1 Autoregressive Moving Average Models (ARMA)

1.1 Characteristic polynomial of AR processes

An AR(p) process y_t is said to be *causal* if it can be written as a one-sided linear process dependent on the past

$$y_t = \Psi(B)\epsilon_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} \tag{1}$$

where B is the backshift operator, with $B\epsilon_t = \epsilon_{t-1}, \psi_0 = 1$ and $\sum_{j=0}^{\infty} |\psi_j| < \infty$.

y_t is causal only when the **autoregressive characteristic polynomial**, defined as

$$\Phi(u) = 1 - \sum_{j=1}^{p} \phi_j u^j \tag{2}$$

has roots with moduli greater than unity. That is, y_t is causal if $\Phi(u) = 0$ only when $|u| > 1$. This causality condition implies stationarity, and so it is often referred as the stationary condition in the time series literature.

The autoregressive characteristic polynomial can also be written as $\Phi(u) = \prod_{j=1}^{p}(1 - \alpha_j u)$, so that its roots are the reciprocals of the α_js. The α_js may be real-valued or may appear as pairs of complex conjugates. Either way, if $|\alpha_j| < 1$ for all j, the process is stationary.

1.2 Structure of ARMA models

Consider a time series y_t, for $t = 1, 2, \cdots$, arising from the model

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \epsilon_t \tag{3}$$

with $\epsilon_t \sim N(0, \nu)$. Then, $\{y_t\}$ follows an autoregressive moving average model, or ARMA(p, q), where p and q are the orders of the autoregressive and moving average parts, respectively. When $p = 0$, $\{y_t\}$ is said to be a moving average process of order q or MA(q). Similarly, when $q = 0$, $\{y_t\}$ is an autoregressive process of order p or AR(p).

Example. MA(1) process. If $\{y_t\}$ follows a MA(1) process, $y_t = \theta\epsilon_{t-1} + \epsilon_t$, the process is stationary for all the values of θ. In addition, it is easy to see that the autocorrelation function has the following form

$$\rho(h) = \begin{cases} 1 & h = 0 \\ \dfrac{\theta}{1+\theta^2} & h = 1 \\ 0 & o.w. \end{cases} \tag{4}$$

Now, if we consider a MA(1) process with coefficient $\frac{1}{\theta}$ instead of θ, we would obtain the same correlation function, and so it would be impossible to determine which of the two processes generated the data. Therefore, it is necessary to impose identifiability conditions on θ. In particular, $\frac{1}{\theta} > 1$ is the identifiability condition for a MA(1), which is also known as the **invertibility condition**, given that it implies that the MA process can be "inverted" into an infinite order AR process.

In general, a MA(q) process is **identifiable or invertible** only when the roots of the **MA characteristic polynomial** $\Theta(u) = 1 + \theta_1 u + \cdots + \theta_q u^q$ lie outside the unit circle. In this case it is possible to write the MA process as an infinite order AR process.

For an ARMA(p, q) process, the stationary condition is given in terms of the AR coefficients, i.e., the process is causal only when the roots of the AR characteristic polynomial $\Phi(u) = 1 - \phi_1 u - \cdots - \phi_p u^p$ lie outside the unit circle. The ARMA process is invertible only when the roots of the MA characteristic polynomial lie outside the unit circle. **So, if the ARMA process is causal and invertible, it can be written either as a purely AR process of infinite order, or as a purely MA process of infinite order.**

If $\{y_t\}$ follows an ARMA(p, q) we can write $\Phi(B)y_t = \Theta(B)\epsilon_t$, with

$$\begin{aligned} \Phi(B) &= 1 - \phi_1 B - \cdots - \phi_p B^p \\ \Theta(B) &= 1 + \theta_1 B + \cdots + \theta_q B^q \end{aligned} \tag{5}$$

where B is the backshift operator. If the process is causal then we can write it as a purely MA process of infinite order

$$y_t = \Phi^{-1}(B)\Theta(B)\epsilon_t = \Psi(B)\epsilon_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j} \tag{6}$$

with $\Psi(B)$ such that $\Phi(B)\Psi(B) = \Theta(B)$. The ψ_j values can be found by solving the homogeneous difference equations given by

$$\psi_j - \sum_{h=1}^{p} \phi_h \psi_{j-h} = 0, \quad j \geq \max(p, q+1) \tag{7}$$

with initial conditions

$$\psi_j - \sum_{h=1}^{j} \phi_h \psi_{j-h} = \theta_j, \quad 0 \leq j < \max(p, q+1) \tag{8}$$

and θ_0. The general solution to the Equations (7) and (8) is given by

$$\psi_j = \alpha_1^j p_1(j) + \cdots + \alpha_r^j p_r(j) \tag{9}$$

where $\alpha_1, \cdots, \alpha_r$ are the reciprocal roots of the characteristic polynomial $\Phi(u) = 0$, with multiplicities $m_1, \cdots, m_r$, respectively, and each $p_i(j)$ is a polynomial of degree $m_i - 1$.

1.3 Inversion of AR components

In contexts where the time series has a reasonable length, we can fit long order AR models rather than ARMA or other, more complex forms. One key reason is that the statistical analysis, at least the conditional analysis based on fixed initial values, is much easier.

If this view is adopted in a given problem, it may be informative to use the results of an AR analysis to explore possible MA component structure using the device of inversion, or partial inversion, of the AR model. Assume that $\{y_t\}$ follows an AR(p) model with parameter vector $\boldsymbol{\phi} = (\phi_1, \cdots, \phi_p)^T$, so we can write

$$\Phi(B)y_t = \prod_{i=1}^{p}(1 - \alpha_i B)y_t = \epsilon_t \tag{10}$$

where the α_is are the autoregressive characteristic reciprocal roots.

For some positive integer $r < p$, suppose that the final $p - r$ reciprocal roots are identified as having moduli less than unity; some or all of the first r roots may also represent stationary components, through that is not necessary for the following development. Then, we can rewrite the model as

$$\prod_{i=1}^{r}(1 - \alpha_i B)y_t = \prod_{i=r+1}^{p}(1 - \alpha_i B)^{-1}\epsilon_t = \Psi^*(B)\epsilon_t \tag{11}$$

where the implicity infinite order MA component has the coefficients of the infinite order polynomial $\Psi^*(u) = 1 + \sum_{j=1}^{\infty} \psi_j^* u^j$, defined by

$$1 = \Psi^*(u) \prod_{i=r+1}^{p}(1 - \alpha_i u) \tag{12}$$

So we have the representation

$$y_t = \sum_{j=1}^{r} \phi_j^* y_{t-j} + \epsilon_t + \sum_{j=1}^{\infty} \psi_j^* \epsilon_{t-j} \tag{13}$$

where the r new AR coefifients ϕ_j^*, for $j = 1, \cdots, r$, are defined by the characteristic equation $\Phi^*(u) = (1 - \alpha_i u) = 0$. The MA terms ψ_j^* can be easily calculated recursively, up to some appropriate upper bound on their number, say q. Explicitly, they are recursively computed as follows.

1. Initialize the algorithm by setting $\psi_i^* = 0$ for all $i = 1 : q$.

2. For $i = (r + 1) : p$, update $\psi_1^* = \psi_1^* + \alpha_i$, and then

   - for $j = 2 : q$, update $\psi_j^* = \psi_j^* + \alpha_i \psi_{j-1}^*$.

Suppose $\boldsymbol{\phi}$ is set at some estimate, such as a posterior mean, in the AR(p) model analysis. The above calculations can be performed for any specified value of r to compute the corresponding MA coefficients in an inversion to the approximating ARMA(r, q) model. If the posterior for $\boldsymbol{\phi}$ is sampled in the AR analysis, the above computations can be performed repeated for all sampled $\boldsymbol{\phi}$ vectors, so producing corresponding samples of the ARMA parameters $\boldsymbol{\phi}^*$ and $\boldsymbol{\psi}^*$. **Thus, inference in various relevant ARMA models can be directly, and quite easily, deduced by inversion of longer order AR models.**

Typically, various values of r will be explored. Guidance is derived from the estimated amplitudes and, in the case of complex roots, periods of the roots of the AR model. Analyses in which some components are persistent suggest that these components should be retained in the AR description. The remaining roots, typically corresponding to high frequency characteristics in the data with lower moduli, are then the candidates for inversion to what will often be a relatively low order MA component. The calculations can be repeated, sequentially increasing q and exploring inferences about the MA parameters, to assess a relevant approximating order.

2 Smoothing and Differencing

Many time series models are built under the stationary assumption. However, in many practical scenarios the data are realizations from one or several nonstationary processes. In this case, methods that aim to eliminate the nonstationary components are often used. The idea is to separate the nonstationary components, such as trends or seasonality, from the stationary ones so that the latter can be carefully studied via traditional time series models such as the aforementioned ARMA models. We briefly discuss two methods that are commonly used in practice for detrending and smoothing.

2.1 Differencing

Differencing is used to remove trends in time series data. The first difference of a time series is defined in terms of the difference operator that we denoted as D, that produces the transformation $Dy_t = y_t - y_{t-1}$. Higher order differences are obtained by successively applying the operator D. For example

$$D^2 y_t = D(Dy_t) = D(y_t - y_{t-1}) = y_t - 2y_{t-1} + y_{t-2} \tag{14}$$

Differencing can also be written in terms of the so called backshift operator B, with $By_t = y_{t-1}$ so that $Dy_t = (1 - B)y_t$ and $D^d y_t = (1 - B)^d y_t$.

2.2 Moving Averages

Moving averages is a method commonly used to "smooth" a time series by removing certain features (e.g., seasonality) to highlight other features (e.g., trends). A moving average is a weighted average of the time series around a particular time t. In general, if we have data $y_{1:T}$, we could obtain a new time series such that

$$z_t = \sum_{j=-q}^{p} a_j y_{t+j} \tag{15}$$

for $t = (q+1) : (T-p)$, with $a_j \geq 0$ and $\sum_{j=-q}^{p} a_j = 1$. Usually $p = q$ and $a_j = a_{-j}$. For example, to remove seasonality in monthly data, one can use a moving average with $p = 6$, $a_6 = a_{-6} = 1/24$, and $a_j = a_{-j} = 1/12$ for $k = 0, \cdots, 5$, resulting in

$$z_t = \frac{1}{24} y_{t-6} + \frac{1}{12} y_{t-5} + \cdots + \frac{1}{12} y_{t+5} + \frac{1}{24} y_{t+6} \tag{16}$$

3 Epilogue

With the methodology we have discussed in this reading material, you should now be confident that the AR models can deal with a large class of time series data. In practice, one can first check the stationarity of the time series. If it contains nonstationary features, try using some detrending, deseasonalizing and smoothing method. Then for the resulting stationary series, using ARMA models and perform the inference by fitting a longer order AR model and inverting AR components.

Reference

Prado, Raquel and West, Mike (2010), "Time Series: Modeling, Computation, and Inference", *Chapman and Hall CRC Texts in Statistical Science*, ISBN 978-1-4200-9336-0.