Lecture 6: Aggregation

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

Aggregation is an interesting subject in time series analysis. Two good references for this topic are the textbook by Wei (1990), which contains various results of aggregation, and the paper by Engel (1984, J. Time Series Analysis, pp. 159-171), which considers models of the sums and aggregations of linear ARMA models.

There are two types of aggregation: sum over series and sum over time (or temporal aggregation). Some relevant results are as follows:

a. Sum of two independent time series: Suppose that X_{1t} and X_{2t} are two independent ARMA series of orders (p_1, q_1) and (p_2, q_2), respectively. Let $Z_t = X_{1t} + X_{2t}$. Then, Z_t is an ARMA(p, q) process with $p \le p_1 + p_2$ and $q \le \max\{p_1 + q_2, p_2 + q_1\}$.

The reason that "$\le$" is used is because of the possibility of common factors in the polynomials involved.

Proof: Write the model for X_{it} as

$$\phi_i(B)X_{it} = \theta_i(B)a_{it}.$$

Applying $\phi_1(B)\phi_2(B)$ to Z_t, we have

$$\phi_1(B)\phi_2(B)Z_t = \phi_1(B)\phi_2(B)(X_{1t} + X_{2t}) = \phi_2(B)\theta_1(B)a_{1t} + \phi_1(B)\theta_2(B)a_{2t}.$$

The result follows.

Notice that the "independent" assumption is important, otherwise, X_{1t} may be a lagged version of X_{2t}, e.g. $X_{1t} = X_{2,t-100}$, which can easily violate the stated result.

b. Temporal aggregation: For a time series $\{Y_t\}$, let $\{Z_\ell\}$ be the series consisting of sums of m non-overlapping points of Y_t. For example, for $m = 3$, we can aggregate monthly series to obtain a quarterly series. Mathematically, we can define Z_ℓ as

$$Z_\ell = \sum_{t=m(\ell-1)+1}^{m\ell} Y_t = (1 + B + \cdots + B^{m-1})Y_{m\ell}.$$

If Y_t is an ARMA(p, q) process, then Z_ℓ is also an ARMA process. To determine the model for Z_t, one needs to consider first the implication of the "backshift operator B" on Z_ℓ. By definition,

$$BZ_\ell = Z_{\ell-1} = \sum_{t=m(\ell-2)+1}^{m(\ell-1)} Y_t = (1 + B + \cdots + B^{m-1})Y_{m(\ell-1)} = B^m(1 + B + \cdots + B^{m-1})Y_{m\ell}.$$

Thus, the backshift operator "B" applying to Z_ℓ is amount to applying B^m to Y_t. In other words, we need to understand that "time scales"are different between the original series Y_t

and the aggregated series Z_t. For instance, a backshift operator in a quarterly time series is corresponding to B^3 in a monthly series.

Example: The best way to derive a model for the aggregated series Z_t is to consider a simple example. Suppose that Y_t is an ARMA(1,1) process, say

$$(1-\phi B)Y_t = (1-\theta B)a_t,$$

and $m = 2$. Applying the polynomial operator $(1-\phi^2 B)$ to Z_ℓ, we have

$$\begin{aligned}
(1-\phi^2 B)Z_\ell &= Z_\ell - \phi^2 Z_{\ell-1} \\
&= (Y_{2\ell} + Y_{2\ell-1}) - \phi^2(Y_{2(\ell-1)} + Y_{2(\ell-1)-1}) \\
&= Y_{2\ell} - \phi^2 Y_{2\ell-2} + Y_{2\ell-1} - \phi^2 Y_{2\ell-3} \\
&= a_{2\ell} + (\phi-\theta)a_{2\ell-1} - \phi\theta a_{2\ell-2} + a_{2\ell-1} + (\phi-\theta)a_{2\ell-2} - \phi\theta a_{2\ell-3} \\
&\overset{def}{=} a_{2\ell} + \omega_1 a_{2\ell-1} + \omega_2 a_{2\ell-2} + \omega_3 a_{2\ell-3}.
\end{aligned}$$

In the above, we have used the result

$$\begin{aligned}
Y_t &= \phi Y_{t-1} + a_t - \theta a_{t-1} \\
&= \phi^2 Y_{t-2} + a_t + (\phi-\theta)a_{t-1} - \phi\theta a_{t-2}.
\end{aligned}$$

Since $2\ell-2$ and $2\ell-3$ correspond to $Z_{\ell-1}$, thus the model for Z_ℓ can be written as

$$(1-\phi^2 B)Z_\ell = (1-\Theta B)b_\ell,$$

which is again an ARMA(1,1) model in the time scale of ℓ.

Example: Consider now that $m = 2$ and Y_t is an AR(2) model, say

$$(1-.5B)(1+.5B)Y_t = (1-.25B^2)Y_t = a_t.$$

Here the coefficients are chosen purposely so that $(.5)^2 = (-.5)^2$. Applying $(1-.25B)$ to Z_ℓ, we have

$$\begin{aligned}
(1-.25B)Z_\ell &= Z_\ell - .25Z_{\ell-1} \\
&= (Y_{2\ell} + Y_{2\ell-1}) - .25(Y_{2\ell-2} + Y_{2\ell-3}) \\
&= (Y_{2\ell} - .25Y_{2\ell-2}) + (Y_{2\ell-1} - .25Y_{2\ell-3}) \\
&= a_{2\ell} + a_{2\ell-1} \\
&\overset{def}{=} b_\ell,
\end{aligned}$$

which is an AR(1) model in the ℓ time scale. Thus, the fact that the coefficients satisfy $(.5)^2 = (-.5)^2$ reduces the AR order by 1. This phenomenon holds in general aggregation.

General result: For an ARMA(p,q) process Y_t, say $\phi(B)Y_t = \theta(B)a_t$, its m-point non-overlapping aggregate Z_t is an ARMA(p^*,q^*) where $p^* \le p$ and $q^* \le [p+1+\frac{q-p-1}{m}]$ with $[x]$ the integer part of x. The inequality "$<$" occurs when some zeros of $\phi(B)$ satisfy $\lambda_i^m = \lambda_j^m$ for $i \ne j$. If all the zeros of $\phi(B)$ are distinct and $\lambda_i^m \ne \lambda_j^m$ for $i \ne j$, then the equality holds. For further details, see Wei (1990, Chap. 16).

Systematic sampling: Suppose that a time series Y_t is an ARMA(p,q) process. However, one only observes the m-th subseries Z_t, which is defined as

$$Z_t = Y_{mt} \quad \text{with} \quad m \quad \text{a fixed positive integer.}$$

This situation occurs frequently in business and in process control. For instance, one may inspect every 5-th product of a production line. Here again, a backshift operator of Z_t is corresponding to B^m of the original series Y_t. In general, the model for Z_t is ARMA(p,r) where $r \le [p+\frac{q-p}{m}]$.