EXERCISES IN FORECASTING III

The Identification of ARMA Models

An appropriate ARMA model for fitting to a stationary series may be identified by inspecting the following thee functions:

(1) The Empirical Autocorrelation Function,
(2) The Empirical Partial Autocorrelation Function,
(3) The Nonparametric Estimate of the Spectral Density Function.

The procedures for determining the orders of an ARMA process from the first two of these functions are spelt out in detail in Lecture 8 of the Yellow Book which is titled *Identification of ARIMA Models*. There are numerous accounts of the so-called Box–Jenkins methodology of model identification which can serve the same purpose. The account in Chapter 3 of Holden, Peel and Thompson *Economic Forecasting* is also worth reading; but it is not entirely adequate and there are no diagrams, which is a serious deficiency.

The Pseudo-Random Data Series

Model identification is essentially a matter of practice. Therefore a large collection of pseudo-random computer-generated data is provided in the files `ARMADATA` and `XYARMA` which you should examine in detail. The processes which have generated the data in `ARMADATA` are revealed in the following list. Those which have generated the data series in `XYARMA` will be revealed to you only after you have attempted to guess the orders of the processes:

(i)
$$\begin{aligned} \text{1ARMA01: } y(t) &= (1-\theta L)\varepsilon(t) \\ y(t) &= (1+0.75L)\varepsilon(t) \end{aligned}$$

(ii)
$$\begin{aligned} \text{2ARMA01: } y(t) &= (1-\theta L)\varepsilon(t) \\ y(t) &= (1-0.75L)\varepsilon(t) \end{aligned}$$

(iii)
$$\begin{aligned} \text{1ARMA10: } (1-\phi L)y(t) &= \varepsilon(t) \\ (1-0.75L)y(t) &= \varepsilon(t) \end{aligned}$$

(iv)
$$\begin{aligned} \text{2ARMA10: } (1-\phi L)y(t) &= \varepsilon(t) \\ (1+0.75L)y(t) &= \varepsilon(t) \end{aligned}$$

(v)
$$\begin{aligned} \text{1ARMA11: } (1-\phi L)y(t) &= (1-\theta L)\varepsilon(t) \\ (1-0.9L)y(t) &= (1+0.9L)\varepsilon(t) \end{aligned}$$

(vi)
$$\begin{aligned} \text{2ARMA11: } (1-\phi L)y(t) &= (1-\theta L)\varepsilon(t) \\ (1+0.9L)y(t) &= (1-0.9L)\varepsilon(t) \end{aligned}$$

(vii)
$$\begin{aligned} \text{1ARMA20: } (1+\alpha_1 L+\alpha_2 L^2)y(t) &= \varepsilon(t) \\ (1-1.273L-0.81L^2)y(t) &= \varepsilon(t) \end{aligned}$$

(viii)
$$\begin{aligned} \text{2ARMA20: } (1+\alpha_1 L+\alpha_2 L^2)y(t) &= \varepsilon(t) \\ (1+1.85L-0.855L^2)y(t) &= \varepsilon(t) \end{aligned}$$

$$\text{(ix)}\qquad \begin{aligned} \text{1ARMA02: } y(t) &= (1 + \mu_1 L + \mu_2 L^2)\varepsilon(t) \\ y(t) &= (1 - 1.273L - 0.81L^2)\varepsilon(t) \end{aligned}$$

$$\text{(x)}\qquad \begin{aligned} \text{2ARMA02: } y(t) &= (1 + \mu_1 L + \mu_2 L^2)\varepsilon(t) \\ y(t) &= (1 + 1.85L - 0.855L^2)\varepsilon(t) \end{aligned}$$

$$\text{(xi)}\qquad \begin{aligned} \text{1ARMA21: } (1 + \alpha L + \alpha_2 L^2)y(t) &= (1 + \mu L)\varepsilon(t) \\ (1 - 1.785L + 0.9025L^2)y(t) &= (1 + 0.95L)\varepsilon(t) \end{aligned}$$

$$\text{(xii)}\qquad \begin{aligned} \text{2ARMA21: } (1 + \alpha L + \alpha_2 L^2)y(t) &= (1 + \mu L)\varepsilon(t) \\ (1 + 1.691L + 0.81L^2)y(t) &= (1 - 0.95L)\varepsilon(t) \end{aligned}$$

$$\text{(xiii)}\qquad \begin{aligned} \text{1ARMA22: } (1 + \alpha L + \alpha_2 L^2)y(t) &= (1 + \mu_1 L + \mu_2 L^2)\varepsilon(t) \\ (1 - 1.4745L + 0.51L^2)y(t) &= (1 - 1.157L + 0.81L^2)\varepsilon(t) \end{aligned}$$

$$\text{(xiv)}\qquad \begin{aligned} \text{2ARMA22: } (1 + \alpha L + \alpha_2 L^2)y(t) &= (1 + \mu_1 L + \mu_2 L^2) \\ (1 - 1.275L + 0.81L^2)y(t) &= (1 + 1.273L + 0.81L^2)\varepsilon(t) \end{aligned}$$

You should plot each of these data series before examining the autocorrelation function and the partial autocorrelation function. You should understand how the features of these functions reflect those of the corresponding theoretical functions which would be generated by the true parameters of the process. The feature of the theoretical function can be attributed to the model orders and to the values of the parameters in the way which is spelt out in Lecture 8.

The Periodogram and the Estimated Spectrum

It is probable that, by the end of the course, the periodogram and the estimated spectral density function or "spectrum" will become your preferred means of model identification. A further explanation of these functions is required before you can use them effectively. Nevertheless, it is appropriate, at this stage, to gather some impression of how their features relate to the parameters of the underlying processes which have generated the data.

Unless a process is very evidently a seasonal or a cyclical one, the raw periodogram is liable to have an irregular or volatile appearance. In order to a provide meaningful estimate of the spectral density function of the process, the profile of the periodogram must be smoothed. This is easily accomplished in *MESOSAUR*. After selecting a variable, you select **Spectrum** from the **Statistics** menu. Then you specify the number of frequency points which you want, and you press `<Enter>`. The default value is 100 points. Finally, you specify a so-called window width. In *MESOSAUR,* the spectral density function is calculated by the Parzen window technique. This involves discarding the higher-order autocovariances and applying a differential weighting scheme to the remainder. The widow width is simply the number of autocovariances which are retained; and, the smaller the width, the smoother is the estimate. It is a remarkable fact that this technique of truncating and weighting the autocorrelation function is mathematically equivalent to smoothing the profile of the periodogram by taking a moving average of its ordinates. The mathematical connection between the autocorrelation fucntion and the periodogram has been established at the end of Lecture 2: *Seasons and Cycles* in the Yellow Book.

1XARMA: MA(2) $y(t) = (1 + 0.75L + 0.5L^2)\varepsilon(t)$

2XARMA: AR(2) $(1 - 1.832L + 0.95L^2)y(t) = \varepsilon(t)$

3XARMA: ARMA(2, 2) $(1 - 1.273L + 0.81L^2)y(t) = (1 + 1.273L + 0.81L^2)\varepsilon(t)$

4XARMA: AR(2) $(1 - 1.724L + 0.7275L^2 y(t) = \varepsilon(t)$

5XARMA: AR(1) $(1 - 0.65L)y(t) = \varepsilon(t)$

6XARMA: AR(4) $(1 - 0.656L^4)y(t) = \varepsilon(t)$

7XARMA: MA(1) $y(t) = (1 + 0.95L)\varepsilon(t)$

8XARMA: MA(2) $y(t) = (1 + L + 0.75L^2)\varepsilon(t)$

9XARMA: AR(3) $(1 - 2.775L + 2.67L^2 - 0.893L^3)y(t) = \varepsilon(t)$

1YARMA: AR(4) $(1 - 0.95L^4)y(t) = \varepsilon(t)$

2YARMA: AR(2) $(1 + 1.871L + 0.9025L^2)y(t) = \varepsilon(t)$

3YARMA: AR(1) $(1 - 0.99L)y(t) = \varepsilon(t)$