In earlier chapters we have considered only univariate models; we now proceed to examine multi-series extensions and to compare the multi-series innovations models with other multi-series schemes. We shall refer to our approach as the vector exponential smoothing (VES) framework. The innovations framework is similar to the structural time series models advocated by Harvey (1989) in that both rely upon unobserved components, but there is a fundamental difference: in keeping with the earlier developments in this book, each time series has only one source of error.

The VES models are introduced in Sect. 17.1; special cases of the general model are then discussed in Sect. 17.2. An inferential framework is then developed in Sect. 17.3 for the VES models, building upon our earlier results for the univariate schemes.

The most commonly used multivariate time series models are those defined within the ARIMA framework. Interestingly, this approach also has only one source of randomness for each time series. Thus, the vector versions of the ARIMA framework (VARIMA), and special cases such as vector autoregression (VAR) and vector moving average (VMA), may be classified as innovations approaches to time series analysis (Lütkepohl 2005). We compare the VES framework with existing approaches in Sect. 17.4. As in Chap. 11, when we consider equivalences between vector innovations models and the VARIMA forms, we will make the infinite start-up assumption.

Finally we compare the performance of VES models to VAR and other existing state space alternatives, first in an empirical study of exchange rates (Sect. 17.5), and then across a range of different time series taken from a large macroeconomic database, in Sect. 17.6.

---

1 This chapter is based on de Silva et al. (2007), which should be consulted for further details.

2 Dr. Ashton de Silva, School of Economics, Finance & Marketing, RMIT University, Melbourne, Australia.
17.1 The Vector Exponential Smoothing Framework

The general vector exponential smoothing model (VES) is introduced in this section. Conceptually, the model builds directly upon the univariate framework; see, for example, Anderson and Moore (1979). We stack the univariate observations into an $N$-vector and then assume that the vector of observations $y_t$ is a linear function of a $k$-vector of unobserved components $x_{t-1}$ plus error. That is, we have the measurement equation

$$y_t = W x_{t-1} + \varepsilon_t,$$

(17.1a)

where $W$ is an $N \times k$ matrix of coefficients that are often known, as in the univariate case, and $\varepsilon_t$ is an $N$-vector. The innovations $\{\varepsilon_t\}$ follow a common multivariate Gaussian distribution with zero means and variance matrix $\Sigma$, but we assume that $\{\varepsilon_t\}$ and $\{\varepsilon_{t+i}\}$ are independent for all $i \neq 0$.

The evolution of the unobserved components is governed by the first-order Markovian relationship

$$x_t = F x_{t-1} + G \varepsilon_t.$$  

(17.1b)

This is called the transition equation. The fixed $k \times k$ matrix $F$ is referred to as the transition matrix. The $k \times N$ matrix $G$ typically has elements that are unknown; they determine the effects of the innovations on the process beyond the period in which they occur. When $G = 0$, the components are deterministic. When $G$ is block-diagonal with some non-zero elements within each block, each innovation has an effect only on its own series. When $G$ has non-zero elements outside these blocks, an innovation will have an effect on other series as well as its own.

The general model given in (17.1) is rather opaque and will often be “parameter-heavy.” A common formulation separates out the state variables for each series, so that the elements of (17.1) may be written as:

$$x_t = \begin{bmatrix} x_{1t} \\ \vdots \\ x_{Nt} \end{bmatrix}, \quad W = \begin{bmatrix} w_1' & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & w_N' \end{bmatrix},$$

$$F = \begin{bmatrix} F_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & F_N \end{bmatrix} \quad \text{and} \quad g = \begin{bmatrix} g_{11} & \cdots & g_{1N} \\ g_{21} & \cdots & g_{2N} \\ \vdots & \ddots & \vdots \\ g_{N1} & \cdots & g_{NN} \end{bmatrix}.$$  

That is, the state variables are updated as functions of the random errors of all series but there are no common states. We now examine several special cases.