A Brief Tutorial

On

Exponential Smoothing Models

Major Observation: Exponential smoothing models are **special cases** of Box-Jenkins models. When adopting exponential smoothing models for forecasting you are implicitly claiming that the same Box-Jenkins model fits all time series equally well. In many instances, exponential smoothing methods may forecast well, but there may also be many instances where building a general Box-Jenkins model for forecasting would do even better. If one only has a few observations to begin with (say 10 or less) then an ad hoc choice of the smoothing parameter (α) (say α = 0.3) can at least generate forecasts whereas having less than 10 observations makes it difficult to use the Box-Jenkins method for building an adequate forecasting model. One has to wait on more observations before progressing from the exponential smoothing models to Box-Jenkins models.

1. The Single Exponential Smoothing (SES) Model is equivalent to a ARIMA(0,1,1) Box-Jenkins

The Single Exponential Smoothing (SES) Model can be written as

$$F_t = \alpha A_{t-1} + (1-\alpha)F_{t-1} \tag{1}$$

where

F_t = exponentially smoothed forecast for period t

A_{t-1} = actual value in prior period

F_{t-1} = exponentially smoothed forecast for period t-1

α = smoothing constant .

If α is close to zero, the forecast relies heavily on past observations (i.e. is smoothed heavily). If $\alpha = 1$ the evidence of past data is ignored completely and the forecast is given by the value of the current observation.

But model (1) is equivalent to the ARIMA(0,1,1) model

$$\Delta A_t = a_t - \theta_1 a_{t-1} \tag{2}$$

where a_t represents a white noise error term and the correspondence between model (1) and model (2) is $\alpha = 1 - \theta_1$. Therefore model (1), the single exponential smoothing model, is a **special case** of the ARIMA(0,1,1) Box-Jenkins model. The above point was made by A.C. Harvey in his book Time Series Models (1981, Philip Allan Publishers), p. 168. This brings up the following Point. Why should one attempt to fit all time series using one particular Box-Jenkins model? In model (2) there is no drift term, therefore the single exponential smoothing model should not be applied to data with trend. A naïve user of the SES model, however, is not likely to know of this subtle point.

One can use MLE on the ARIMA(0,1,1) model of (2) and obtain an estimate of the smoothing parameter α as $\hat{\alpha} = 1 - \hat{\theta}_1$, where $\hat{\theta}_1$ is the MLE of the moving average order-one parameter θ_1. Alternatively, one could do a grid search over the interval [0,1] and apply (1) to determine the RMSEs of one-step forecasts for different values of α and choose the α value that minimizes RMSE. (This is essentially the same as using MLE to estimate the ARIMA(0,1,1) model (2)).

2. The Seasonal Single Exponential Smoothing (SSES) Model is equivalent to a ARIMA(0,0,0)x(0,1,1)s Multiplicative Seasonal Box-Jenkins model.

The seasonal single exponential smoothing (SSES) model is of the form

$$F_t = \alpha A_{t-s} + (1-\alpha) F_{t-s} \quad . \tag{3}$$

This model is equivalent to the ARIMA(0,0,0)x(0,1,1) Multiplicative Seasonal Box-Jenkins Model

$$\Delta_s A_t = a_t - \Theta_1 a_{t-s} \quad . \tag{4}$$

We have the correspondence $\alpha = 1 - \Theta_1$. Therefore, one can estimate α as $\hat{\alpha} = 1 - \hat{\Theta}_1$, where $\hat{\Theta}_1$ is the MLE of Θ_1 in model (4) above.

Again, the question arises as to why one would restrict the search for a good time series model to a **special case** of the Multiplicative Seasonal Box-Jenkins model. Are all seasonal time series equally well characterized by the same special case Box-Jenkins model? Notice (4) does not allow for trend in the data after taking the seasonal span difference which certainly doesn't apply to all time series data. In many time series the more appropriate differencing of seasonal time series is $\Delta_1 \Delta_s$. One can now appreciate the inflexibility built into the SSES model.