Dynamic Linear Models and the Kalman Filter

Introduction

Dynamic linear models (DLM) are a class of models for time series, including multivariate time series. They are particularly popular among Bayesians although there is no particular reason why Bayesians could not use autoregressive integrated moving average (ARIMA) models, for example.

The idea is that the observation y_t (which may be a vector) at time t depends on an underlying unobserved *state vector* (or *system vector*) β_t. The state vector changes randomly over time and the dependence between y_t and y_s, where $s \neq t$, is modelled only by the dependence between β_t and β_s.

This is a Gaussian process where the joint distribution of $\ldots \beta_{t-2}, \beta_{t-1}, \beta_t, \beta_{t+1}, \beta_{t+2}, \ldots, y_{t-2}, y_{t-1}, y_t, y_{t+1}, y_{t+2} \ldots$ is multivariate normal. Of course people have invented extensions to this where assumptions, for examp[le of normality, are relaxed but we will not concern ourselves with these here.

We have an observation vector at time t :

$$Y_t = X_t \beta_t + \varepsilon_t$$

where X_t is a known matrix and ε_t is a vector of "errors" with mean 0 and (known) variance matrix P_t. The vector Y_t could, for example, be the sales of several products in month t.

We have an underlying state vector

$$\beta_t = T_t \beta_{t-1} + u_t$$

where u_t is a random vector with mean 0 and (known) variance matrix Q_t.

Associated with a DLM there is an algorithm called a *Kalman filter* which allows us to update our beliefs about the current value of the state vector each time we make a new observation.

The "generation step"

At time t we can calculate a "prior" mean and variance for the quantities at time t.

The expectation of β_t at time t is b_t so the expectation of β_t at time $t-1$ is $b_{t|t-1} = T_t b_{t-1}$. We can not observe the value of the state vector but at any time we will have a mean vector and a variance matrix for it.

The variance of β_t at time t is S_t so the variance of β_t at time $t-1$ is $S_{t|t-1} = T_t S_{t-1} T_t' + Q_t$.

At time $t-1$ the expectation of Y_t is

$$F_t = X_t b_{t|t-1},$$

the variance of Y_t is

$$D_t = X_t S_{t|t-1} X_t' + P_t$$

and the covariance of β_t and Y_t is

$$C_t = S_{t|t-1} X_t'.$$

So, at time $t-1$,

$$\mathrm{E}\begin{pmatrix} \beta_t \\ Y_t \end{pmatrix} = \begin{pmatrix} b_{t|t-1} \\ F_t \end{pmatrix}$$

and

$$\mathrm{var}\begin{pmatrix} \beta_t \\ Y_t \end{pmatrix} = \begin{pmatrix} S_{t|t-1} & C_t \\ C_t' & D_t \end{pmatrix}.$$

The "observation step"

At time t we observe Y_t. This means that we update our beliefs about β_t. (If we are assuming that everything is normally distributed then formally this is done by applying Bayes' rule. If we are using Bayes linear methods then we use Bayes linear updating. The formulae and results are the same in both cases).

The updated mean for β_t is

$$b_t = b_{t|t-1} + C_t D_t^{-1}(Y_t - F_t)$$

and the updated variance matrix for β_t is

$$S_t = S_{t|t-1} - C_t D_t^{-1} C_t'.$$

Note that the variance matrices P_t and Q_t are known/given. Although I have put t subscripts on P, Q, X, T, often these would remain constant.

Updating

When we observe some new data, e.g. a new month's sales figure, we carry out first a generation step then an observation step to update our beliefs about the system vector.

Forecasting

A generation step on its own gives a one-step-ahead forecast. We can generate forecasts further into the future by a sequence of generation steps without observation steps. For example, suppose we have observed the data at time t. We can find the one-step-ahead forecasts. (Here I am dropping the t subscripts on P, Q, X, T).

$$\mathrm{E}\begin{pmatrix} \beta_{t+1} \\ Y_{t+1} \end{pmatrix} = \begin{pmatrix} Tb_t \\ XTb_t \end{pmatrix}$$

and

$$\mathrm{var}\begin{pmatrix} \beta_{t+1} \\ Y_{t+1} \end{pmatrix} = \begin{pmatrix} S_{t+1|t} & C_{t+1} \\ C_{t+1}' & D_{t+1} \end{pmatrix}.$$

Now we can calculate the two-step-ahead forecasts.

$$\mathrm{E}\begin{pmatrix} \beta_{t+2} \\ Y_{t+2} \end{pmatrix} = \begin{pmatrix} T & 0 \\ XT & 0 \end{pmatrix}\begin{pmatrix} Tb_t \\ XTb_t \end{pmatrix} = \begin{pmatrix} TTb_t \\ XTTb_t \end{pmatrix}$$

and

$$\mathrm{var}\begin{pmatrix} \beta_{t+2} \\ Y_{t+2} \end{pmatrix} = \begin{pmatrix} T & 0 \\ XT & 0 \end{pmatrix}\begin{pmatrix} S_{t+1|t} & C_{t+1} \\ C_{t+1}' & D_{t+1} \end{pmatrix}\begin{pmatrix} T' & T'X' \\ 0 & 0 \end{pmatrix}.$$

Then we can calculate the three-step-ahead forecasts and so on.

Trend, seasonals etc.

Trend, seasonals etc. are modelled by special forms of the system vector and the matrices T and Q.

Further reading

See West, M. and Harrison, J., 1997, *Bayesian Forecasting and Dynamic Models* (2nd ed.), New York: Springer-Verlag.