Discussion

Markus Hürzeler  
AICOS Technologies AG  
Efringerstrasse 32  
CH-4057 Basel, Switzerland  
mhuerzeler@aicos.com

I would like to thank the three authors for their contribution to the topic of state-space models. They show the wide range of this field concerning different concepts of methods and the variety of applications ranging from technical to economical time series. Fahrmeier gives a comprehensive introduction and quotes the recent developments which concentrate on non-linear and non-Gaussian state-space models. There exists a large class of different methods to compute the easily written down formulas for prediction, filtering and smoothing. Except for some special state-space models all of these methods are approximations of the corresponding formula.

The MCMC approach also used in Fronk and Fahrmeier (1998) and the particle filter of Pitt and Shephard (1999) represent two large and different classes of those methods, where the later keeps the recursive structure of the general Kalman filter recursion and for that reason is well applicable in on-line situations. Whereas the MCMC approach is only suitable in off-line situations. In the following I will focus my discussion on some particular aspects.

Fahrmeier mentions the usage of local smoothing parameters for modelling the stylized functions Blocks, Bumps, HeavySine and Doppler. Especially for the Blocks and HeavySine example there is an unsmooth tendency of the Kalman smoother estimation in smooth areas of the functions, see Fronk and Fahrmeier (1998). This lies in the smoothness assumption on the local variance, i.e., if a jump occurs a large variance is needed and so also in the neighbourhood of this point a larger variance will be used. In my view a more appropriate way to model jumps is to replace the normal density by a heavy-tailed one, i.e., to model the jumps as innovation outliers. The application of this models to the stylized functions gives with the direct smoothing method of Hürzeler and Künsch (1998) a curve estimation which lies near the ideal wavelet method. The ideal wavelet method uses the knowledge of the underlying function whereas the general state-space models does not, see Hürzeler (1998). Since heavy-tailed distributions can be written as mixture of normals with different variances our method with the direct particle filter is implicitly an approach with local varying variances but which are independent between different time points.

Pitt introduces a new variation of a particle filter sampler. Based on the filter sample \((\alpha_t^i)_{i=1,...,M}\) the next filter density can be approximated by

\[
\hat{f}(\alpha_{t+1} | F_{t+1}) \approx \sum_{i=1}^{M} f(y_{t+1} | \alpha_{t+1}) f(\alpha_{t+1} | \alpha_t^i).
\]

Pitt and Shephard (1999) introduce now an additional approximation of formula (1) in order to sample more efficient from this new approximation. The adapted reweighing should also reduce the problems of a few distinguished values which can occur in some applications when we reweigh the prediction sample according to the likelihood, see Kitagawa and Gersch (1996). But also this new approximation proposed by Pitt will lead to few distinguished values in the switching mean model, since the likelihood is strongly peaked and its main support lies in the main region of the prediction density. A direct simulation according to (1) without any resampling was proposed by Hürzeler and Künsch (1998). It would be interesting to compare these two methods in accuracy and computational effort. In my view the computational effort has to be measured in computing time and not only in the order of the sample size \(N\), e.g. the rejection algorithm has an order of \(O(N)\). But
the constants in this order $O(N)$ can be relevant. The direct filter is based on a rejection algorithm and we had no real problems in computing the filter samples in our applications. Moreover the direct filter gives prediction samples and the likelihood without additional effort. This leads to an easy feasible way to compute estimations of hyper-parameters. Even in a full Bayesian framework some hyper-parameters of the assumed time evolution of the parameters remains.

Instead of a general approach to reduce computer costs in the direct sampling according to (1) I would propose to take the specific general state-space model and to look if with these densities a more efficient way for sampling according to (1) is available in those cases where the general direct sampling is not feasible.

The contribution of Kim goes in the direction of implementing the general Kalman recursion for specific models. He introduces a modified Kalman filter algorithm for a state space model with Markov-switching variance. Model II can be seen as a Gaussian Mixture model and so several linear Kalman filter steps have to be made. Since the number of densities in the mixture for the filter density explodes a reduction of the components in the mixture is needed. It would be interesting to look at the effect of different strategies for the mixture reduction and to compare this approach with particle filters.

Finally I hope that in the near future some literature as Doucet et. al (2000) which collect many proposals appear and a lot of these proposed methods become available implemented in some common statistical software packages. As I have learned as statistical consultant in industry the availability is the first criterion for the end-user. This would also encourage all of us to make more comparisons of methods. In addition, I always propose in practice to use for a specific problem two different methods since two similar results provide more confidence in them and from very different solutions we can learn probably more about the underlying problem.

**ADDITIONAL REFERENCES**

