A P P E N D I X

5

Kalman filtering

K. Friston and W. Penny

INTRODUCTION

Bayesian inversion of state-space models is related to Bayesian belief update procedures (i.e. recursive Bayesian filters). The conventional approach to online Bayesian tracking of states in non-linear or non-Gaussian systems employs extended Kalman filtering or sequential Monte Carlo methods, such as particle filtering. These implementations of Bayesian filters approximate the conditional densities of hidden states in a recursive and computationally expedient fashion, assuming that the parameters and hyperparameters of the system are known. We start with systems (dynamic models) that are formulated in continuous time:

$$\begin{aligned} y &= g(x) + z \\ \dot{x} &= f(x, v) \end{aligned} \qquad \textbf{A5.1}$$

where the innovations $z(t)$ and causes $v(t)$ are treated as random fluctuations. As we will see below this is converted into a state-space model in discrete time before application of the filter. Kalman filters proceed recursively in two steps: prediction and update. The prediction uses the Chapman-Kolmogorov equation to compute the density of the hidden states $x(t)$ conditioned on the response up to, but not including, the current observation $y_{\to t-1}$:

$$p(x_t|y_{\to t-1}) = \int p(x_t|x_{t-1})p(x_{t-1}|y_{\to t-1})dx_{t-1} \qquad \textbf{A5.2}$$

This conditional density is then treated as a prior for the next observation and Bayes' rule is used to compute the conditional density of the states, conditioned upon all observations $y_{\to t}$. This gives the Bayesian update:

$$q(x_t) = p(x_t|y_{\to t}) \propto p(y_t|x_t)p(x_t|y_{\to t-1}) \qquad \textbf{A5.3}$$

Critically, the conditional density covers only the hidden states. This is important because it precludes inference on causes and the ability to de-convolve inputs from outputs. This is a key limitation of Bayesian filtering. However, Kalman filtering provides the optimal solution when the assumptions of the underlying model hold and one is not interested in causes or inputs. We now consider in more detail the operation equations for the extended Kalman filter. The extended Kalman filter is a generalization of the Kalman filter, in which the linear operators of the state equations are replaced by the partial derivatives of $f(x, v)$ with respect to the states.

THE EXTENDED KALMAN FILTER

This section provides a pseudo-code specification of the extended Kalman filter based on van der Merwe *et al.* (2000). To clarify the notation, we will use $f_x = \partial f/\partial x$. Eqn. **A5.1** can be re-written, using local linearization, as a discrete-time state-space model. This is the formulation treated in Bayesian filtering procedures:

$$\begin{aligned} y_t &= \mathbf{g}_x x_t + z_t \\ x_t &= \mathbf{f}_x x_{t-1} + w_{t-1} \\ \mathbf{g}_x &= g(x_t)_x \\ \mathbf{f}_x &= \exp(f(x_t)_x) \\ z_t &= z(t) \\ w_{t-1} &= \int \exp(f_x\tau)f_v v(t-\tau)d\tau \end{aligned} \qquad \textbf{A5.4}$$

For simplicity, we assume $\Delta t = 1$. The key thing to note here is that process noise w_{t-1} is simply a convolution of

Statistical Parametric Mapping, by Karl Friston et al.
ISBN–13: 978-0-12-372560-8 ISBN–10: 0-12-372560-7

Copyright 2007, Elsevier Ltd. All rights reserved.

the causes $v(t)$. This is relevant for Kalman filtering and related non-linear Bayesian tracking schemes that assume w_{t-1} is a well-behaved noise sequence. The covariance of process noise is:

$$\begin{aligned}\langle w_t w_t^T \rangle &= \int \exp(f_x \tau)\Omega \exp(f_x \tau)^T d\tau \\ &\approx \Omega \\ &= f_v R f_v^T\end{aligned} \qquad \textbf{A5.5}$$

where R is the covariance of $v(t)$. We have assumed $v(t)$ has no temporal correlations and that the Lyapunov exponents of f_x are large relative to the time-step. The prediction and update steps are:

for all t

Prediction step

$$x_t = \mathbf{f}_\mathbf{x} x_{t-1}$$
$$\Sigma_t^x = \Omega + \mathbf{f}_x \Sigma_{t-1}^x \mathbf{f}_x^T$$

Update or correction step

$$K = \Sigma_t^x g_x^T (\Sigma + g_x \Sigma_t^x g_x^T)^{-1}$$
$$x_t \leftarrow x_t + K(y - g(x_t))$$
$$\Sigma_t^x \leftarrow (I - K g_x)\Sigma_t^x$$

end **A5.6**

Where Σ is the covariance of observation noise. The Kalman gain matrix K is used to update the prediction of future states and their conditional covariance, given each new observation. We actually use $x_t = x_{t-1} + (f_x - I)$ $f_x^{-1} f(x_{t-1})$. This is a slightly more sophisticated update that uses the current state as the expansion point for the local linearization. As mentioned in Chapter 37, Kalman filtering is also known as variable parameter regression, when the hidden state plays the role of a parameter (see Büchel and Friston, 1998).

REFERENCES

Büchel C, Friston KJ (1998) Dynamic changes in effective connectivity characterised by variable parameter regression and Kalman filtering. *Hum Brain Mapp* **6**: 403–08

van der Merwe R, Doucet A, de Freitas N *et al.* (2000) The unscented particle filter. Technical Report CUED/F-INFENG/TR 380