Adaptive ensemble Kalman filtering of nonlinear systems

Tyrus Berry

January 15, 2014

Nonlinear Kalman-type Filter: Problem Setup

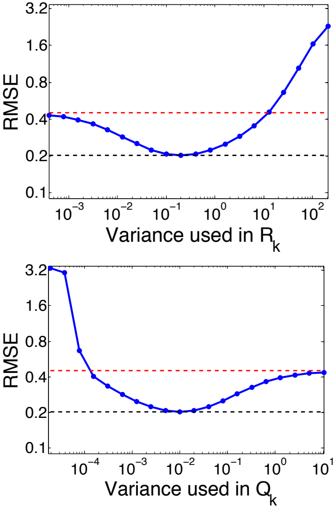
- We consider a system of the form:

$$
\begin{aligned}
x_{k+1} &= f(x_k) + \omega_{k+1} \qquad && \omega \approx \mathcal{N}(0, Q) \\
y_{k+1} &= h(x_{k+1}) + \nu_{k+1} \qquad && \nu \approx \mathcal{N}(0, R)
\end{aligned}
$$

- We initially assume Gaussian system and observation noise
- Our goal is to estimate the covariance matrices Q and R as part of the filter procedure
- Later we consider Q to be an additive inflation which attempts to compensate for model error

Nonlinear Kalman-type Filter: Influence of Q and R

- Simple example with full observation and diagonal noise covariances
- Red indicates RMSE of unfiltered observations
- Black is RMSE of 'optimal' filter (true covariances known)



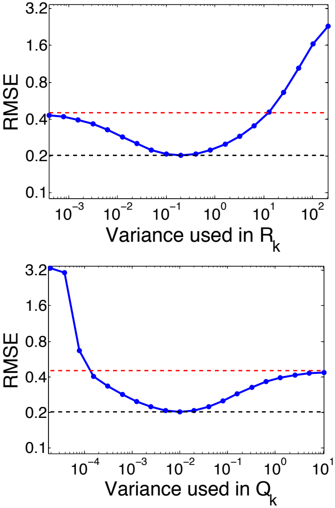
Nonlinear Kalman-type Filter: Influence of Q and R

Standard Kalman Update:

$$
\begin{aligned}
P_k^f &= F_{k-1} P_{k-1}^a F_{k-1}^T + Q_{k-1} \\
P_k^y &= H_k P_k^f H_k^T + R_{k-1}
\end{aligned}
$$

$$
\begin{aligned}
K_k &= P_k^f H_k^T (P_k^y)^{-1} \\
P_k^a &= (I - K_k H_k) P_k^f
\end{aligned}
$$

$$
\begin{aligned}
\epsilon_k &= y_k - y_k^f = y_k - H_k x_k^f \\
x_k^a &= x_k^f + K_k \epsilon_k
\end{aligned}
$$



Adaptive Filter: Estimating Q and R

- Innovations contain information about Q and R

$$\begin{aligned}
\epsilon_k &= y_k - y_k^f \\
&= h(x_k) + \nu_k - h(x_k^f) \\
&= h(f(x_{k-1}) + \omega_k) - h(f(x_{k-1}^a)) + \nu_k \\
&\approx H_k F_{k-1}(x_{k-1} - x_{k-1}^a) + H_k\omega_k + \nu_k
\end{aligned}$$

- IDEA: Use innovations to produce samples of Q and R :

$$\begin{aligned}
\mathbb{E}[\epsilon_k \epsilon_k^T] &\approx HP^f H^T + R \\
\mathbb{E}[\epsilon_{k+1} \epsilon_k^T] &\approx HFP^e H^T - HFK\mathbb{E}[\epsilon_k \epsilon_k^T] \\
P^e &\approx FP^a F^T + Q
\end{aligned}$$

- In the linear case this is rigorous and was first solved by Mehra in 1970

Adaptive Filter: Estimating Q and R

- To find Q and R we estimate H_k and F_{k-1} from the ensemble and invert the equations:

$$\begin{aligned} \mathbb{E}[\epsilon_k \epsilon_k^T] &\approx HP^f H^T + R \\ \mathbb{E}[\epsilon_{k+1} \epsilon_k^T] &\approx HFP^e H^T - HFK\mathbb{E}[\epsilon_k \epsilon_k^T] \end{aligned}$$

- This gives the following *empirical* estimates of Q_k and R_k:

$$\begin{aligned} P_k^e &= (H_{k+1}F_k)^{-1}(\epsilon_{k+1}\epsilon_k^T + H_{k+1}F_k K_k \epsilon_k \epsilon_k^T)H_k^{-T} \\ Q_k^e &= P_k^e - F_{k-1}P_{k-1}^a F_{k-1}^T \\ R_k^e &= \epsilon_k \epsilon_k^T - H_k P_k^f H_k^T \end{aligned}$$

- Note: P_k^e is an empirical estimate of the background covariance

An Adaptive Kalman-Type Filter for Nonlinear Problems

We combine the estimates of Q and R with a moving average

Original Kalman Eqs.

$$P_k^f = F_{k-1} P_{k-1}^a F_{k-1}^T + Q_{k-1}$$
$$P_k^y = H_k P_k^f H_k^T + R_{k-1}$$

$$K_k = P_k^f H_k^T (P_k^y)^{-1}$$
$$P_k^a = (I - K_k H_k) P_k^f$$

$$\epsilon_k = y_k - y_k^f$$
$$x_k^a = x_k^f + K_k \epsilon_k$$

Our Additional Update

$$P_{k-1}^e = F_{k-1}^{-1} H_k^{-1} \epsilon_k \epsilon_{k-1}^T H_{k-1}^{-T} + K_{k-1} \epsilon_{k-1} \epsilon_{k-1}^T H_{k-1}^{-T}$$

$$Q_{k-1}^e = P_{k-1}^e - F_{k-2} P_{k-2}^a F_{k-2}^T$$
$$R_{k-1}^e = \epsilon_{k-1} \epsilon_{k-1}^T - H_{k-1} P_{k-1}^f H_{k-1}^T$$

$$Q_k = Q_{k-1} + (Q_{k-1}^e - Q_{k-1})/\tau$$
$$R_k = R_{k-1} + (R_{k-1}^e - R_{k-1})/\tau$$

How does this compare to inflation?

- We extend Kalman's equations to estimate Q and R
- Estimates converge for linear models with Gaussian noise
- When applied to nonlinear, non-Gaussian problems
  - We interpret Q as an additive inflation
  - Q can have complex structure, possibly more effective than multiplicative inflation?
  - Downside: many more parameters than multiplicative inflation
- Somewhat less ad hoc than other inflation techniques?

Observability and Parameterization of Q

Recall:

$$\begin{aligned} P^e_{k-1} &= F^{-1}_{k-1} H^{-1}_k \epsilon_k \epsilon^T_{k-1} H^{-T}_{k-1} + K_{k-1}\epsilon_{k-1}\epsilon^T_{k-1} H^{-T}_{k-1} \\ Q^e_{k-1} &= P^e_{k-1} - F_{k-2} P^a_{k-2} F^T_{k-2} \end{aligned}$$

Together these equations imply that:

$$\begin{aligned} H_k F_{k-1} Q^e_k H^T_{k-1} &= \epsilon_k \epsilon^T_{k-1} + H_k F_{k-1} K_{k-1} \epsilon_{k-1} \epsilon^T_{k-1} \\ &\quad - H_k F_{k-1} P^a_{k-1} F^T_{k-1} H^T_{k-1} \end{aligned}$$

Set C_k equal to the right hand side (we simply compute C_k). Parameterize $Q^e_k = \sum_{i=1}^s q_i \hat{Q}_i$ where q_i are scalar parameters and $\hat{Q}_i$ are 'shape' matrices.

Observability and Parameterization of Q

We now need to solve:

$$C_k = \sum_{i=1}^{s} q_i H_k F_{k-1} \hat{Q}_i H_{k-1}^T$$

We vectorize the equation as

$$\text{vec}(C_k) = \sum_{i=1}^{s} q_i \text{vec}(H_k F_{k-1} \hat{Q}_i H_{k-1}^T) = A_k [q_1, ..., q_s]^T$$

where A_k is an m^2-by-l matrix where the i-th row is given by $\text{vec}(H_k F_{k-1} \hat{Q}_i H_{k-1}^T)$.
We can the solve for the parameters $[q_1, ..., q_s]^T$ by least squares.

Adaptive Filter: Application to Lorenz-96

- We will apply the adaptive EnKF to the 40-dimensional Lorenz96 model integrated over a time step $\Delta t = 0.05$

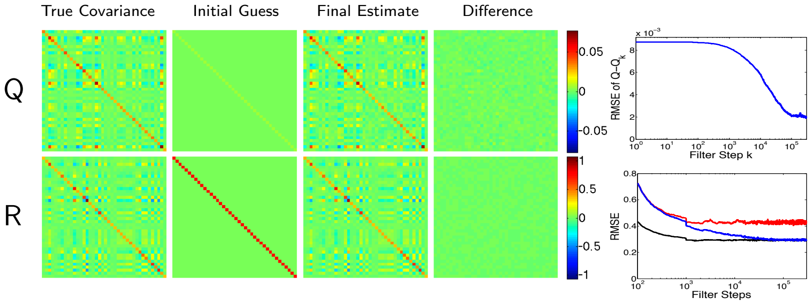
$$\frac{dx^i}{dt} = -x^{i-2}x^{i-1} + x^{i-1}x^{i+1} - x^i + F$$

- We augment the model with Gaussian white noise

$$\begin{aligned} x_k &= f(x_{k-1}) + \omega_k & \qquad \omega_k = \mathcal{N}(0, Q) \\ y_k &= h(x_k) + \nu_k & \qquad \nu_k = \mathcal{N}(0, R) \end{aligned}$$

- We will consider full and sparse observations
- The Adaptive EnKF uses $F = 8$
- We will consider model error where the true $F^i = \mathcal{N}(8, 16)$

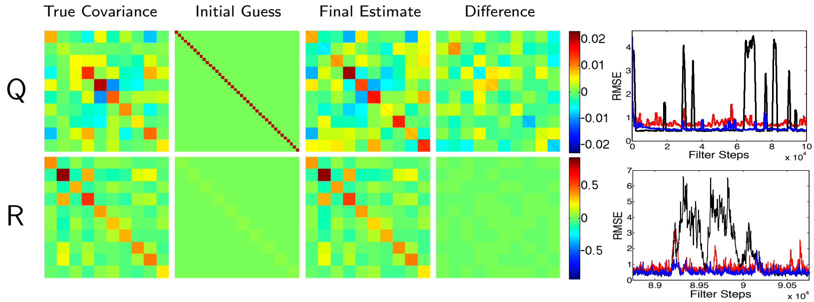
Recovering Q and R, Full Observability



RMSE shown for the initial guess covariances (red) the true Q and R (black) and the adaptive filter (blue)

Recovering Q and R, Sparse Observability

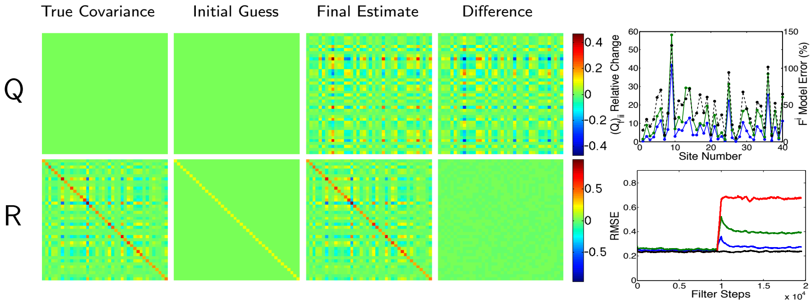
Observing 10 sites results in divergence with the true Q and R



RMSE shown for the initial guess covariances (red) the true Q and R (black) and the adaptive filter (blue)

Compensating for Model Error

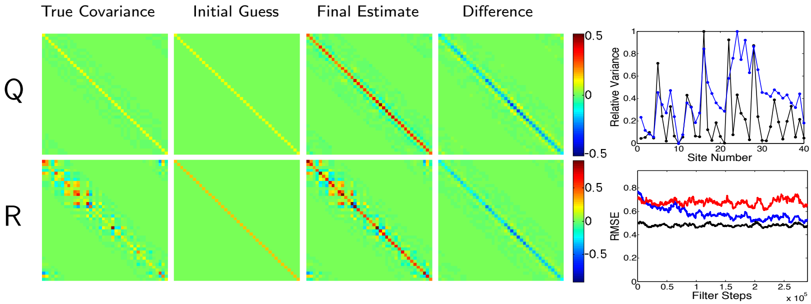
The adaptive filter compensates for errors in the forcing F^i



RMSE shown for the initial guess covariances (red) an Oracle EnKF (black) and the adaptive filter (blue)

Integration with the LETKF

Simply find a local Q and R for each region



RMSE shown for the initial guess covariances (red) the true Q and R (black) and the adaptive filter (blue)

This research was partially supported by the National Science Foundation Directorates of Engineering (EFRI-1024713) and Mathematics and Physical Sciences (DMS-1216568).

- T. Berry, T. Sauer, Adaptive ensemble Kalman filtering of nonlinear systems. To appear in Tellus A.
- R. Mehra, 1970: On the identification of variances and adaptive Kalman filtering.
- P. R. Bélanger, 1974: Estimation of noise covariance matrices for a linear time-varying stochastic process.
- J. Anderson, 2007: An adaptive covariance inflation error correction algorithm for ensemble filters.
- H. Li, E. Kalnay, T. Miyoshi, 2009: Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter.
- B. Hunt, E. Kostelich, I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter.
- E. Ott, et al. 2004: A local ensemble Kalman filter for atmospheric data assimilation.