Adaptive ensemble Kalman filtering of nonlinear systems

Tyrus Berry and Tim Sauer

Department of Mathematical Sciences
George Mason University
Fairfax VA 22030

Overview

A necessary ingredient of an ensemble Kalman filter is covariance inflation, used to control filter divergence and compensate for model error. There is an ongoing search for inflation tunings that can be learned adaptively. Early in the development of Kalman filtering, Mehra enabled adaptivity in the context of linear dynamics with white noise model errors by showing how to estimate the model error and observation covariances. We propose an adaptive scheme, based on lifting Mehra's idea to the nonlinear case, that recovers the model error and observation noise covariances in simple cases, and in more complicated cases, finds a natural additive inflation that improves state estimation.

Adaptive Scheme for Kalman Filtering

In its most general form our adaptive scheme can be interpreted as an extension of the Kalman update (red) which updates the matrices Q_k and R_k using the filter innovation ϵ_k as shown below (blue). Let $x_k^f = F_{k-1}x_{k-1}^a$ be the state forecast and $y_k^f = H_k x_k^f$ the observation forecast. The innovation $\epsilon_k = y_k - y_k^f$ defines the following updates

$$\begin{aligned} P_k^f &= F_{k-1}P_{k-1}^a F_{k-1}^T + Q_{k-1} & P_{k-1}^e &= F_{k-1}^{-1}H_k^{-1}(\epsilon_k\epsilon_{k-1}^T + H_kF_{k-1}K_{k-1}\epsilon_{k-1}\epsilon_{k-1}^T)H_{k-1}^{-T} \\ P_k^y &= H_kP_k^fH_k^T + R_{k-1} & Q_{k-1}^e &= P_{k-1}^e - F_{k-2}P_{k-2}^aF_{k-2}^T \\ K_k &= P_k^fH_k^T(P_k^y)^{-1} & R_{k-1}^e &= \epsilon_{k-1}\epsilon_{k-1}^T - H_{k-1}P_{k-1}^fH_{k-1}^T \\ P_k^a &= (I - K_kH_k)P_k^f & Q_k &= Q_{k-1} + (Q_{k-1}^e - Q_{k-1})/\tau \\ x_k^a &= x_k^f + K_k\epsilon_k & R_k &= R_{k-1} + (R_{k-1}^e - R_{k-1})/\tau \end{aligned}$$

where τ is called the stationarity parameter and controls the speed at which the Q_k and R_k estimates adapt. In the examples we show that when the model error and observation noise are given by Gaussian white noise our adaptive scheme recovers the covariances of these distributions as Q_k and R_k. For more realistic types of model error we interpret Q_k as an adaptive inflation and demonstrate improvements in state estimation. We will demonstrate the adaptive scheme in the context of the ensemble Kalman filter (EnKF) and the local ensemble transform Kalman filter (LETKF).

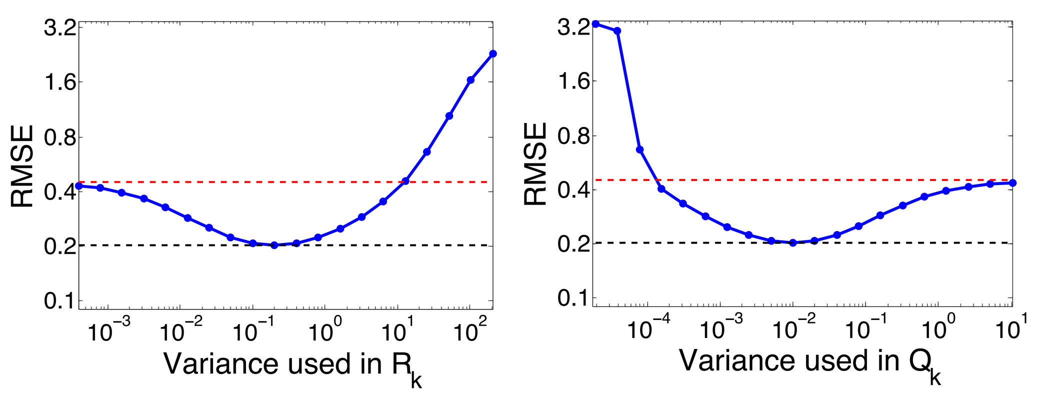
Application to Lorenz96

We will apply the adaptive EnKF to the 40-dimensional Lorenz96 model

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

where $x = [x^1(t), \ldots, x^{40}(t)] \in \mathbb{R}^{40}$ and the superscript refers to the ith vector coordinate. We work in the discrete setting by defining $f(x_{k-1})$ to be the result of integrating the above system for $\Delta t = 0.05$ with initial condition x_{k-1}. Unless otherwise stated we will use $F = 8$ and the full observation $h(x_k) = x_k$. We augment the model with Gaussian white noise

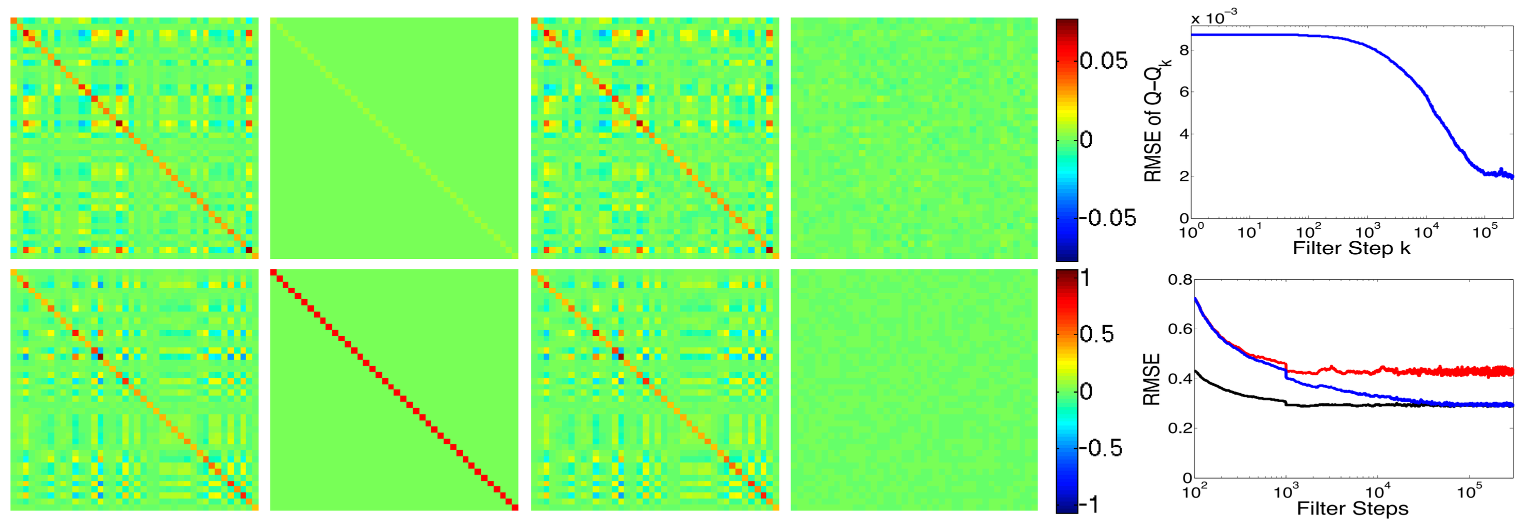
$$\begin{aligned} x_k &= f(x_{k-1}) + \omega_k & \omega_k &= \mathcal{N}(0, Q) \\ y_k &= h(x_k) + \nu_k & \nu_k &= \mathcal{N}(0, R). \end{aligned}$$



Above we illustrate the effect of using sub-optimal filters by running a conventional EnKF on data simulated from the Lorenz96 model with $Q = (0.01)I_{40}$ and $R = (0.2)I_{40}$. The RMS error of the signal prior to filtering was $\sqrt{0.2} \approx 0.45$ (red dotted line) the RMSE of the optimal filter using $Q_k = Q$ and $R_k = R$ was 0.20 (black dotted line). We show the effect of varying R_k when $Q_k = Q$ and the effect of varying Q_k when $R_k = R$. We see that in one extreme the filter becomes trivial, and in the other extreme it is possible for the filter to actually degrade the signal. This shows the importance of Q_k and R_k to filter performance.

Gaussian White Noise Covariances

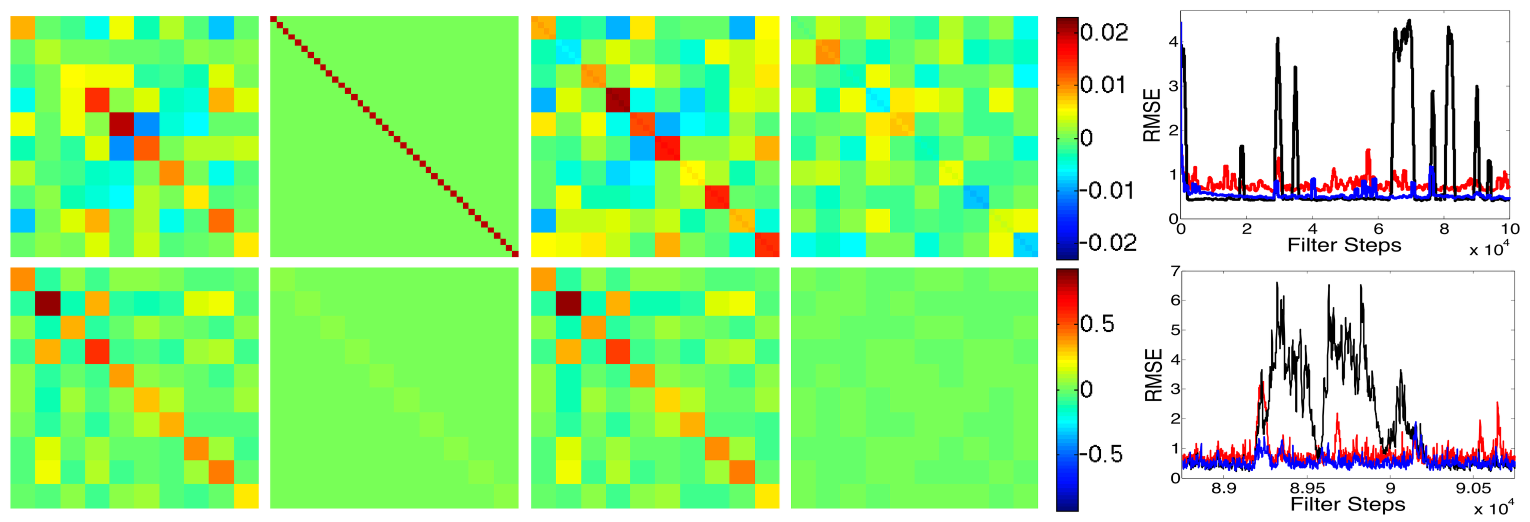
In this example we show the long term performance of the adaptive EnKF by simulating Lorenz96 for 300000 steps with Q and R matrices that were randomly generated symmetric positive definite matrices (shown below, first column). We then initialized Q_k and R_k as diagonal matrices (below, second column). The adaptive EnKF was then applied to the simulated data with stationarity $\tau = 20000$ and the final estimates of Q_k and R_k are shown in the third column along with the final differences $Q - Q_k$ and $R - R_k$ in the forth column.



In the top right plot we show how the RMS difference between the entries of Q and Q_k declines as the adaptive EnKF recovers the covariance structure. The limiting RMS difference can be decreased by increasing the stationarity τ, however this increases the number of steps needed for the adaptive EnKF to converge. In the bottom right plot we show the RMS error of the state estimates for adaptive EnKF (blue) compared to the conventional EnKF run with the true Q and R (black) and with the diagonal guess matrices (red).

Sparse Observations with Gaussian Noise

In this example we apply the adaptive EnKF to a low-dimensional observation. In particular we use a sparse observation which is given by observing 10 equally spaced sites among the 40 total sites. Since H_{k-1} or H_kF_{k-1} will not be invertible for rank deficient observations, we parameterize $Q_k^e = \sum q_pQ_p$ as a linear combination of fixed matrices Q_p. For this example we choose a block constant parameterization with 100 blocks of size 4×4. Due to symmetry there are 55 parameters which allows us to solve for Q_k^e at each step.

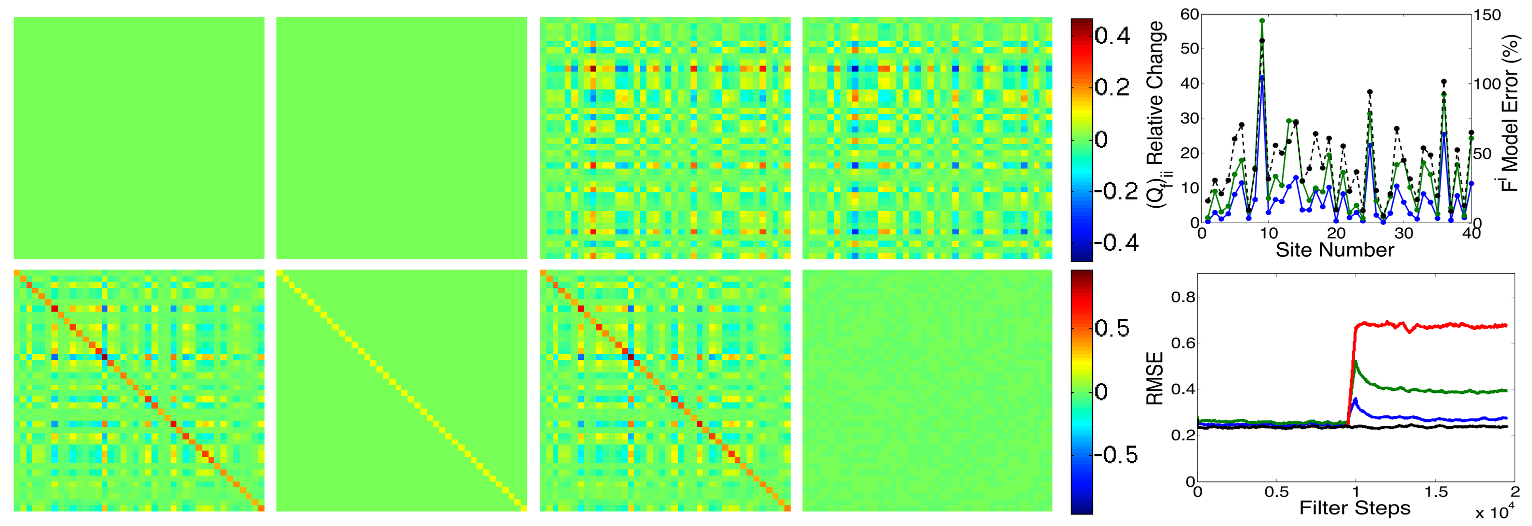


To illustrate the inflation found by the adaptive EnKF we generated a random Q matrix with the block constant structure and a random R matrix (above, first column). The Lorenz96 model was simulated for 100000 steps with these covariances. We chose diagonal guess matrices (second column) and ran the adaptive EnKF with the block constant parameterization of Q_k and $\tau = 15000$. The final estimates of Q_k and R_k are shown above in the third column along with the final differences $Q - Q_k$ and $R - R_k$. In the fifth column we show the RMSE of the state estimates for the adaptive EnKF (blue) compared to the conventional EnKF run with the true Q and R (black) and the diagonal guess matrices (red).

Such a low dimensional observation dramatically increases the RMSE of the state estimate, as shown above. Moreover, we now observe filter divergence, where the state estimate trajectory completely loses track of the true trajectory. Filter divergence occurs only when the true matrix Q is used in the filter, whereas both the initial guess and the final estimate produced by the adaptive EnKF are *inflated*. This example shows how the breakdown of the assumptions of the EnKF (local linearizations, Gaussian distributions) can lead to model error even when the nonlinear dynamics are known. In the presence of this model error, our adaptive EnKF must be interpreted as an inflation scheme and we judge it by its performance in terms of RMSE rather than recovery of the underlying Q.

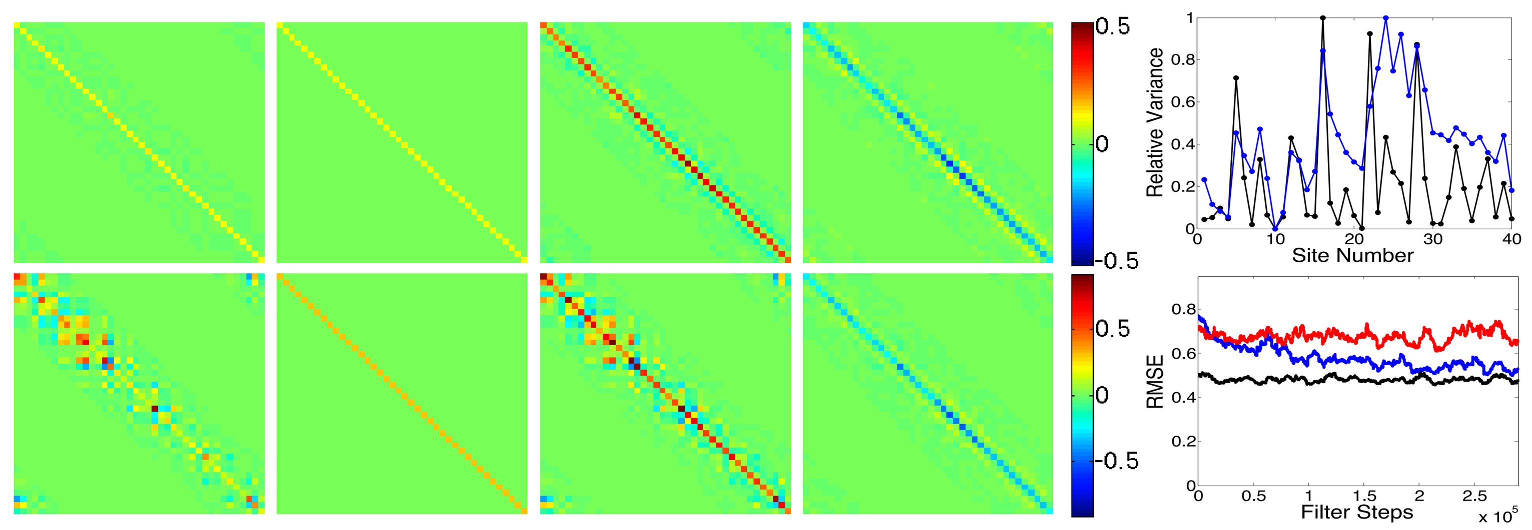
Compensating for Model Error

This example shows that the covariance structure Q_k can compensate for systematic model error. We simulated Lorenz96 with full observation for 10000 steps with $F = 8$ and then chose 40 fixed random values of $F^i = \mathcal{N}(8, 16)$ and continued the simulation for another 10000 steps. The model used by the filters had $F = 8$ fixed. The matrices below show the true Q and R, the diagonal guess, the final estimates, and the final differences, as in previous examples. In the presence of the model error the adaptive EnKF still recovers R and in the top right figure we show that the inflation variances found in the final Q_k (blue) correlated with the amount of model error (black) at each site. The adaptive EnKF significantly reduced the RMSE of the state estimate (blue) compared to that of a conventional EnKF with the true Q and R matrices (red). The RMSE for an oracle EnKF (given the parameters F^i) is in black, and the adaptive EnKF with Q_k diagonal is in green.



An Adaptive LETKF

This example shows that the adaptive scheme can naturally integrate into the local ensemble transform Kalman filter (LETKF). The LETKF uses the spatial structure of the state space to perform the Kalman update locally. We used the algorithm of Ott et al. and performed the local Kalman update at each site using a local region with 11 sites ($l = 5$) and a global ensemble with 22 members. Since the LETKF uses a Kalman update for each local region, we simply estimated 40 separate 11×11 local matrices Q_k^i and R_k^i using our adaptive scheme. For display purposes we integrated the final estimates into a global versions Q_k and R_k shown in the third column. The adaptive LETKF (blue) significantly improved the RMSE compared to simply using a diagonal inflation (red).



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

References

[1] J. Anderson, 2007: An adaptive covariance inflation error correction algorithm for ensemble filters.
[2] P. R. Bélanger, 1974: Estimation of noise covariance matrices for a linear time-varying stochastic process.
[3] B. Hunt, E. Kostelich, I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter.
[4] B. Hunt, E. Kalnay, E. Kostelich, E. Ott, D.J. Patil, T. Sauer, I. Szunyogh, J. Yorke, A. Zimin, 2004: Four-dimensional ensemble Kalman filtering.
[5] R. Kalman, 1960: A new approach to linear filtering and prediction problems.
[6] E. Kalnay, 2003: Atmospheric modeling, data assimilation, and predictability.
[7] H. Li, E. Kalnay, T. Miyoshi, 2009: Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter.
[8] E. Lorenz, 1996: Predictability: A problem partly solved.
[9] R. Mehra, 1970: On the identification of variances and adaptive Kalman filtering.
[10] E. Ott, et al. 2004: A local ensemble Kalman filter for atmospheric data assimilation.