Stability of the Ensemble Kalman Filter

David Kelly Andy Majda Xin Tong

Courant Institute
New York University
New York NY

dtbkelly.com

July 8, 2015

Equadiff 2015
Lyon.

Talk overview

The Ensemble Kalman Filter (EnKF) is a data assimilation algorithm used for very high dimensional nonlinear models.

It is an ‘approximation’ of the Kalman filter.

EnKF inherits stability properties from the underlying model.

The filtering problem

We have a **model** (deterministic, for now)

$$\frac{dv}{dt} = F(v) \quad \text{with } v_0 \sim N(m_0, C_0)\ .$$

We will denote $v(t) = \Psi_t(v_0)$. Think of this as **very high dimensional** and **nonlinear**.

We want to **estimate** $v_n = v(nh)$ for some $h > 0$ and $n = 0, 1, 2, \ldots$ given the **observations**

$$y_n = Hv_n + \xi_n \quad \text{for } \xi_n \text{ iid } N(0, \Gamma).$$

In the linear setting, the Kalman filter gives an exact expression for the posterior $\mathbf{P}(v_{n+1}|y_{n+1}, v_n)$

EnKF **approximates** this procedure in two ways: first the posterior is represented **empirically** via samples and second, the samples are **not actually samples**.

Sampling from the posterior with linear model

Suppose we are given K samples $\{u_n^{(1)}, \dots, u_n^{(K)}\}$ from the time n posterior. Here is how we turn them into samples from the $n+1$ posterior.

For each ensemble member (sample), we create an **artificial observation**

$$y_{n+1}^{(k)} = y_{n+1} + \xi_{n+1}^{(k)} \quad , \quad \xi_{n+1}^{(k)} \text{ iid } N(0, \Gamma).$$

We update each member using the **Kalman update**

$$u_{n+1}^{(k)} = \Psi_h(u_n^{(k)}) + G_n \left(y_{n+1}^{(k)} - H\Psi_h(u_n^{(k)}) \right) ,$$

where G_n is the **Kalman gain matrix** .

The EnKF approximation

Suppose we are 'approximate samples' $\{u_n^{(1)}, \ldots, u_n^{(K)}\}$ from the time n posterior. For each ensemble member, we create an **artificial observation**

$$y_{n+1}^{(k)} = y_{n+1} + \xi_{n+1}^{(k)} \quad , \quad \xi_{n+1}^{(k)} \text{ iid } N(0, \Gamma).$$

We update each member using the **Kalman update**

$$u_{n+1}^{(k)} = \Psi_h(u_n^{(k)}) + G(u_n)\left(y_{n+1}^{(k)} - H\Psi_h(u_n^{(k)})\right) \; ,$$

where $G(u_n)$ is the **Kalman gain** computed using the **forecasted ensemble covariance**

$$\widehat{C}_{n+1} = \frac{1}{K}\sum_{k=1}^{K}(\Psi_h(u_n^{(k)}) - \overline{\Psi_h(u_n)})^T(\Psi_h(u_n^{(k)}) - \overline{\Psi_h(u_n)}) \; .$$

Stability properties of EnKF

Stability #1: Model 'dissipativity' is inherited by the filter.

Assumptions on the dynamics

The state v satisfies an **energy (dissipation) criterion**:

$$\mathbf{E}_n|v_{n+1}|^2 - |v_n|^2 \leq -\beta|v_n|^2 + K$$

for some $\beta \in (0,1)$ and $K > 0$. $\mathbf{E}_n$ is expectation conditioned on everything up to time n.

Eg. The finite dimensional SDE

$$\frac{dv}{dt} + Av + B(v,v) = f$$

with A linear elliptic, B is an energy preserving bilinearity, f is stochastic forcing.

Nb. Infinite dimensions are possible, but not done here.

Assumptions on the observations

The observation matrix H must be chosen in such a way that

$$\mathbf{E}_n|Hv_{n+1}|^2 - |Hv_n|^2 \leq -\beta|Hv_n|^2 + K$$

We call this the **observable energy criterion**.

ie. If there is an effective subspace controlling the dynamics then H observes this subspace.

Eg. $v = (v^{(1)}, v^{(2)})$ where $v^{(1)}$ are **slow** variables and $v^{(2)}$ are **fast** variables. Suppose that $Hv = v^{(1)}$. The slow variables can be approximated by an effective system $\frac{d\overline{v}}{dt} = F(\overline{v})$ which is dissipative.

Theorem (Tong, Majda, K. '15)

The EnKF satisfies the energy criterion

$$\mathbf{E}_n\left(\mathcal{E}_{n+1}\right) - \mathcal{E}_n \leq -\beta' \mathcal{E}_n + K'$$

where $\mathcal{E}_n = |Hv_n|^2 + \sum_{k=1}^K \lambda |Hu_n^{(k)}|$ *and* $\beta' \in (0,1)$, $K' > 0$.

Consequently, the observed components of EnKF are bounded (in mean square sense) uniformly in time:

$$\sup_{n\geq 1} \sum_{k=1}^{K} \mathbf{E}|Hu_n^{(k)}|^2 < \infty \; .$$

Rmk 1. The bound may seem trivial, but EnKF is known (numerically) to explode to machine infinity, for very turbulent models (Harlim, Majda '11 & Gottwald, Majda '13).

Rmk 2. Improvement on (K, Law, Stuart '14) which shows at most exponential growth in the fully observed case.

Proving it

From the update equation for EnKF

$$Hu_{n+1}^{(k)} = (I + H\widehat{C}_{n+1}H^T)^{-1}H\Psi_h(u_n^{(k)}) + H\widehat{C}_nH^T(I + H\widehat{C}_{n+1}H^T)^{-1}y_{n+1}^{(k)}$$

In calculating $\mathbf{E}_n|u_{n+1}^{(k)}|^2$, the first term is controlled using the observable energy criterion and the second term is controlled using the observable energy criterion + finite variance of the noise.

Stability #2: Ergodicity of the model is inherited by the filter.

Assumptions for ergodicity

Assumption 1 - The model-ensemble process $(v, u^{(1)}, \dots, u^{(K)})$ has a Lyapnuov function $\mathcal{E}$ with compact sublevel sets.

Assumption 2 - The noise in the model is non-degenerate and has a density wrt Lebesgue.

Eg. If H is full rank and the model is the SDE

$$dv = b(v)dt + \sigma dW$$

with $b(u) \cdot u \leq -\alpha |u|^2 + c$ and σ full rank.

Theorem (Tong, Majda, K. 15)

The model-ensemble process $(v_n, u_n^{(1)}, \ldots, u_n^{(K)})$ *is geometrically ergodic.*

ie. Let $P^n\mu$ *be the law of* $(v_n, u_n^{(1)}, \ldots, u_n^{(K)})$ *initialized with* $(v_0, u_0^{(1)}, \ldots, u_0^{(K)}) \sim \mu$*, then there exists an unique probability measure* π *with*

$$|P^n\mu - \pi|_{TV} \leq C\gamma^n$$

for some $\gamma \in (0, 1)$.

How does it work?

We use the **Meyn-Tweedie** strategy: Lyapunov function + minorization condition implies geometric ergodicity.

The Lyapunov function is an assumption for us. Sufficient to check the minorization condition.

For a Markov chain X_n, with Kernel P, the **minorization** condition boils down to checking the following: There exists a compact set C such that:

1 - There is an '**intermediate point**' $y^* \in C$ such that for every $\delta > 0, x \in C$ we have $P(x, B_\delta(y^*)) > 0$.
2 - The Markov kernel has a jointly continuous **density** wrt Lebesgue in a nbhd of y^*.

Minorization for EnKF

Recall that $u_{n+1}^{(k)} = \Psi_h(u_n^{(k)}) + G(u_n)\left(y_{n+1}^{(k)} - H\Psi_h(u_n^{(k)})\right)$

The Markov kernel for $(v_n, u_n^{(1)}, \dots, u_n^{(K)})$ can be written $P(x, A) = Q(x, \Gamma^{-1}(A))$ where $Q(x, \cdot)$ is a nice Markov kernel and Γ is a nice function.

$Q(x, \cdot)$ is described by the random mapping

$$(v_n, u_n^{(1)}, \dots, u_n^{(K)}) \mapsto (\Psi_h(v_n), \Psi_h(u_n^{(1)}), \dots, \Psi_h(u_n^{(K)}), y_{n+1}^{(1)}, \dots, y_{n+1}^{(K)})$$

and Γ by

$$(\Psi_h(v_n), \Psi_h(u_n^{(1)}), \dots, \Psi_h(u_n^{(K)}), z_{n+1}^{(1)}, \dots, z_{n+1}^{(K)}) \mapsto (v_{n+1}, u_{n+1}^{(1)}, \dots, u_{n+1}^{(K)})$$

Remarks and coming attractions

For EnKF, Ergodicity requires a Lyapunov function with compact sublevel sets. On the face of it, this requires full rank H.

It is easy to tweak EnKF, via an adaptive inflation, so that it a Lyapnuov function with compact sublevel sets for arbitrary H. Joint work with Majda, Tong. To appear on my website soon.

When does EnKF get it wrong? What causes (catastrophic) filter divergence?

We have built an extremely simple dissipative model for which EnKF exhibits arbitrary long spells of exponential growth, for generic filter initializations. Joint work with Majda, Tong. To appear on my website soon.

Thank you!

Nonlinear stability and ergodicity of ensemble based Kalman filters.
X. Tong, A. Majda, D. Kelly. (2015).

www.dtbkelly.com