Ergodicity in data assimilation methods

David Kelly

Andy Majda Xin Tong

Courant Institute
New York University
New York NY
www.dtbkelly.com

April 15, 2016

AIMS 16, *Orlando, Florida*

Data assimilation and EnKF

Suppose we have a stochastic dynamical system $u_{n+1} = \Psi(u_n)$ with an random initial condition $u_0 \sim N(m_0, C_0)$.

We make noisy, partial observations of the state $y_{n+1} = Hu_{n+1} + \xi_{n+1}$ (H low rank matrix and ξ_n iid Gaussians) and write $Y_n = \{y_1, \dots, y_n\}$.

The Ensemble Kalman Filter is an empirical (but inconsistent) approximation of the posterior

$$\mathbf{P}(u_n | Y_n) \approx \frac{1}{K} \sum_{k=1}^{K} \delta(u_n - u_n^{(k)})$$

EnKF derivation

To derive the EnKF update equations $\{u_n^{(k)}\}_{k=1}^K \mapsto \{u_{n+1}^{(k)}\}_{k=1}^K$, first write down Bayes formula:

$$\mathbf{P}\left(u_{n+1} | Y_{n+1}\right) \propto \mathbf{P}\left(y_{n+1} | u_{n+1}\right) \mathbf{P}\left(u_{n+1} | Y_n\right)$$
$$= \exp\left(-\frac{1}{2}|y_{n+1} - Hu_{n+1}|_\Gamma^2\right) \mathbf{P}\left(u_{n+1} | Y_n\right)$$

Use $\{\Psi^{(k)}(u_n^{(k)})\}_{k=1}^K$ to approximate $\mathbf{P}(u_{n+1}|Y_n)$ with a Gaussian

$$\exp\left(-\frac{1}{2}|y_{n+1} - Hu_{n+1}|_\Gamma^2\right) \exp\left(-\frac{1}{2}|u_{n+1} - \widehat{m}_{n+1}|_{\widehat{C}_{n+1}}^2\right)$$

Draw samples $\{u_{n+1}^{(k)}\}_{k=1}^K$ from resulting Gaussian.

The update equations for EnKF

The EnKF 'particles' $\{u_n^{(k)}\}_{k=1}^K$ are updated according to

$$u_{n+1}^{(k)} = (1 - K_{n+1}H)\Psi^{(k)}(u_n^{(k)}) + K_{n+1}Hy_{n+1}^{(k)}$$

where $\Psi^{(k)}$ are independent realizations of the **dynamics**, $y_{n+1}^{(k)} = y_{n+1} + \xi_{n+1}^{(k)}$ are **perturbed observations** and K_{n+1} is the **empirical Kalman gain** matrix

$$K_{n+1} = \widehat{C}_{n+1}H^T(H\widehat{C}_{n+1}H^T + \Gamma)^{-1} \quad \text{and}$$

$$\widehat{C}_{n+1} = \frac{1}{K-1}\sum_{k=1}^{K}\left(\Psi^{(k)}(u_n^{(k)}) - \overline{\Psi^{(\cdot)}(u_n^{(\cdot)})}\right)\left(\Psi^{(k)}(u_n^{(k)}) - \overline{\Psi^{(\cdot)}(u_n^{(\cdot)})}\right)^T$$

Ergodicity for EnKF

Two notions of ergodicity

1$-$ Signal-filter ergodicity:

Statistics of $(u_n, u_n^{(1)}, \dots, u_n^{(K)})$ *converge to invariant statistics as* $n \to \infty$.

2$-$ Conditional ergodicity:

Laws of $(u_n^{(1)}, \dots, u_n^{(K)}) | Y_n$ *with different intializations* $(u_0^{(1)}, \dots, u_0^{(K)})$ *converge to eachother (and hopefully the posterior) as* $n \to \infty$.

Rmk. All results are in the K fixed regime.

Animation

We have the two-dimensional model

$$du = -\nabla V(u)dt + \sigma dW$$

where $V(x,y) = (1 - x^2 - y^2)^2$ and we only observe the x variable.

EnKF is **signal-filter ergodic**, as the marginals converge to uniform measure on circle. But also **conditionally ergodic**, the law is close to the posterior, regardless of initialization.

Today we focus on geometric ergodicity for the **signal-filter** process $(u_n, u_n^{(1)}, \dots, u_n^{(K)})$.

The theoretical framework

A Markov chain $\{X_n\}_{n\in\mathbb{N}}$ on a state space $\mathcal{X}$ is called **geometrically ergodic** if it has a unique invariant measure π and for any initialization x_0 we have

$$\sup_{|f|\leq 1}\left|\mathbf{E}_{x_0} f(X_n) - \int f(x)\pi(dx)\right| \leq C(x_0) r^n$$

for some $r \in (0,1)$ and any m-ble bdd f.

We use the **coupling** approach: Let X'_n and X''_n be two copies of X_n, such that $X'_0 = x_0$ and $X''_0 \sim \pi$, and are coupled in such a way that $X'_n = X''_n$ for $n \geq T$, where T is the first hitting time $X'_T = X''_T$. Then we have

$$\|P^n \delta_{x_0} - \pi\|_{TV} \leq 2\mathbf{P}(T > n)$$

The **Doeblin** / **Meyn-Tweedie** approach is to verify two assumptions that guarantee the coupling can be constructed with $\mathbf{P}(T > n) \lesssim r^n$:

1- Lyapunov function / Energy dissipation: $\mathbf{E}_n|X_{n+1}|^2 \leq \alpha|X_n|^2 + \beta$ with $\alpha \in (0,1)$, $\beta > 0$.

2- Minorization: Find compact $C \subset \mathcal{X}$, measure ν supported on C, $\kappa > 0$ such that $P(x, A) \geq \kappa\nu(A)$ for all $x \in C$, $A \subset \mathcal{X}$.

To construct the coupling (within C) we let $\tilde{X}_{n+1} = \tilde{f}(\tilde{X}_n, \omega)$ describe the Markov chain with kernel $\tilde{P}(x, A) = \frac{1}{1-\kappa}(P(x, A) - \kappa\nu(A))$ and let

$$X'_{n+1} = \phi\tilde{f}(X'_n, \omega) + (1-\phi)\xi$$

where $\phi \sim \mathrm{Bernoulli}(\kappa)$ and $\xi \sim \nu$. Easy to see that $\mathbf{P}_x(X'_1 \in A) = P(x, A)$, so this is a copy of the original Markov chain.

Lyapunov function: Inheriting an energy principle

Suppose we know the model satisfies an energy principle

$$\mathbf{E}_n|\Psi(u)|^2 \leq \alpha|u|^2 + \beta$$

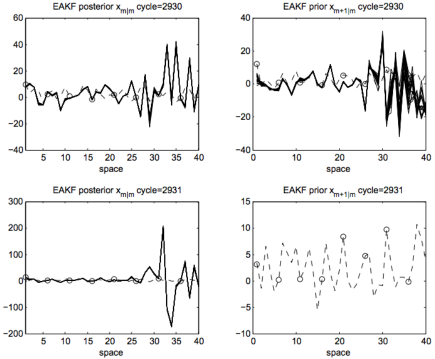
for $\alpha \in (0,1), \beta > 0$. Does the filter inherit the energy principle?

$$\mathbf{E}_n|u_{n+1}^{(k)}|^2 \leq \alpha'|u_n^{(k)}|^2 + \beta'$$

The is strong evidence to suggest that this does **not** hold in general

Catastrophic filter divergence

Lorenz-96: $\dot{u}_j = (u_{j+1} - u_{j-2})u_{j-1} - u_j + F$ with $j = 1, \ldots, 40$. Periodic BCs. Observe every fifth node. *(Harlim-Majda 10, Gottwald-Majda 12)*



True solution in a bounded set, but filter **blows up** to machine infinity in finite time!

The rotate-and-lock map (K., Majda, Tong. PNAS 15.)

The model $\Psi : \mathbb{R}^2 \to \mathbb{R}^2$ is a composition of two maps $\Psi(x, y) = \Psi_{lock}(\Psi_{rot}(x, y))$ where

$$\Psi_{rot}(x, y) = \begin{pmatrix} \rho\cos\theta & -\rho\sin\theta \\ \rho\sin\theta & \rho\cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

and Ψ_{lock} rounds the input to the nearest point in the grid

$$\mathcal{G} = \{(m, (2n+1)\varepsilon) \in \mathbb{R}^2 : m, n \in \mathbb{Z}\} .$$

It is easy to show that this model has an **energy dissipation principle**:

$$|\Psi(x, y)|^2 \le \alpha|(x, y)|^2 + \beta$$

for $\alpha \in (0, 1)$ and $\beta > 0$.

(a)

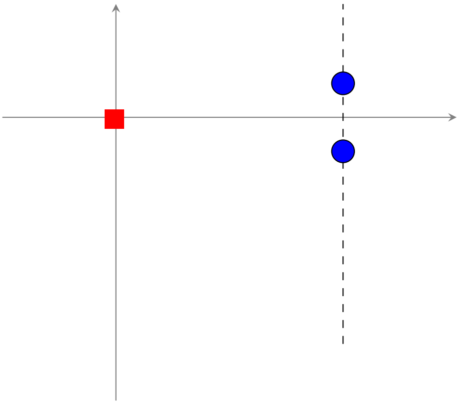


Figure: The red square is the trajectory $u_n = 0$. The blue dots are the positions of the forecast ensemble $\Psi(u_0^+)$, $\Psi(u_0^-)$. Given the locking mechanism in Ψ, this is a natural configuration.

(b)

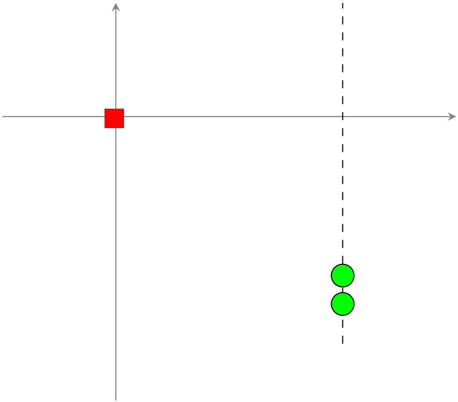


Figure: We make an observation (H shown below) and perform the analysis step. The green dots are u_1^+, u_1^-.

Observation matrix

$$H = \begin{bmatrix} 1 & 0 \\ \varepsilon^{-2} & 1 \end{bmatrix}$$

Truth $u_n = (0,0)$.

The filter is certain that the x-coordinate is $\hat{x}$ (the dashed line). The filter thinks the observation must be $(\hat{x}, \varepsilon^{-2}\hat{x} + u_{1,y})$, but it is actually $(0,0)$. The filter concludes that $u_{1,y} \approx -\varepsilon^{-2}\hat{x}$.

(c)

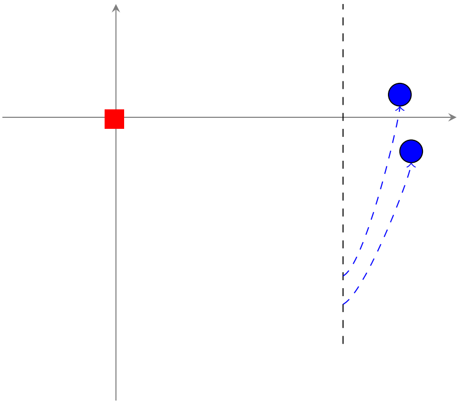


Figure: Beginning the next assimilation step. Apply Ψ_{rot} to the ensemble (blue dots)

(d)

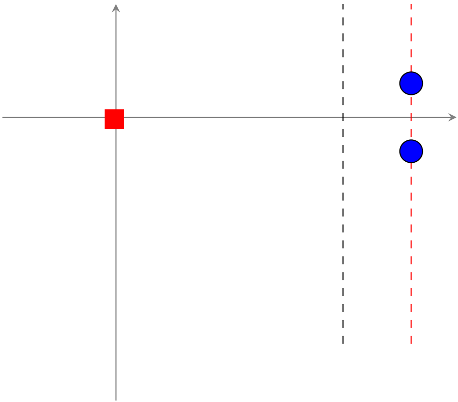


Figure: Apply Ψ_{lock}. The blue dots are the forecast ensemble $\Psi(u_1^+)$, $\Psi(u_1^-)$. Exact same as frame 1, but higher energy orbit. The cycle repeats leading to **exponential growth**.

For any $N > 0$ and any $p \in (0,1)$ there exists a choice of parameters such that

$$\mathbf{P}\left(|u_n^{(k)}| \geq M_n \text{ for all } n \leq N\right) \geq 1 - p$$

where M_n is an exponentially growing sequence.

ie - The filter can be made to grow exponentially for an arbitrarily long time with an arbitrarily high probability.

Ensemble alignment can cause EnKF to gain energy, eventually leading to finite time **blow-up**

This is known as **catastrophic filter divergence**.

Can we get around the problem by **tweaking** the algorithm?

Adaptive Covariance Inflation (Tong, Majda, K. 15)

We modify algorithm by introducing a **covariance inflation** :

$$\widehat{C}_{n+1} \mapsto \widehat{C}_{n+1} + \lambda_{n+1} I$$

where

$$\lambda_{n+1} \propto \Theta_{n+1} \mathbf{1}(\Theta_{n+1} > \Lambda)$$

$$\Theta_{n+1} = \sqrt{\frac{1}{K} \sum_{k=1}^{K} |y_{n+1}^{(k)} - H\Psi(u_n^{(k)})|^2}$$

and Λ is some constant threshold. If the predictions are near the observations, then there is no inflation.

Thm. The modified EnKF inherits an energy principle from the model.

$$\mathbf{E}_x |\Psi(x)|^2 \leq \alpha |x|^2 + \beta \Rightarrow \mathbf{E}_n |u_{n+1}^{(k)}|^2 \leq \alpha' |u_n^{(k)}|^2 + \beta'$$

Consequently, the modified EnKF is signal-filter ergodic.

References

1 - D. Kelly, K. Law & A. Stuart. *Well-Posedness And Accuracy Of The Ensemble Kalman Filter In Discrete And Continuous Time.* **Nonlinearity** (2014).

2 - D. Kelly, A. Majda & X. Tong. *Concrete ensemble Kalman filters with rigorous catastrophic filter divergence*. **Proc. Nat. Acad. Sci.** (2015).

3 - X. Tong, A. Majda & D. Kelly. *Nonlinear stability and ergodicity of ensemble based Kalman filters*. **Nonlinearity** (2016).

4 - X. Tong, A. Majda & D. Kelly. *Nonlinear stability of the ensemble Kalman filter with adaptive covariance inflation*. To appear in **Comm. Math. Sci.** (2016).

All my slides are on my website (www.dtbkelly.com) **Thank you**!