Kalman Filtering and Smoothing for an Advection Equation with Model Error

Andrew M Stuart

University of Warwick
Mathematatcis Institute and
Centre for Scientific Computing

collaboration with

Wonjung Lee, University of Oxford
Damon McDougall, University of Warwick

funded by

EPSRC, ERC, ONR, NERC

advection equation on a torus

We study an advection equation on a torus

$$\partial_t v(x,t) + c \cdot \nabla v(x,t) = 0, \quad (x,t) \in \mathbf{T}^2 \times (0,\infty) \qquad (1)$$
$$\frac{dv}{dt}(t) + \mathcal{L}v(t) = 0,, \quad t \in (0,\infty).$$

- $v(\cdot,0) = v_0 = N(\widehat{m}_0, \widehat{\mathcal{C}}_0)$
- $v(x,t) = v_0(x - ct, 0)$ solves eq. (1)
- the random field $v(\cdot,t)$ is Gaussian

data assimilation

Suppose we have data at every $t_n = n \times \Delta t$

$$y(x, t_n) = v(x, t_n) + \eta(x, t_n), \quad \text{or}$$
$$y_n = v_n + \eta_n \quad \text{where} \quad \eta_n \sim \mathcal{N}(0, \Gamma)$$

then $\mathbb{P}\big(v_n | Y_n = \{y_1, \cdots, y_n\}\big)$ is obtained using Bayes rule

$$\frac{\mathbb{P}(v_n | Y_n)}{\mathbb{P}(v_n | Y_{n-1})} \propto \mathbb{P}(y_n | v_n)$$

- Gaussianity preserved when conditioned on data Y_n
- Infinite dimensional Kalman filter and smoother
- Filter $\mathbb{P}(v_n | Y_n) = N(\widehat{m}_n, \widehat{C}_n)$
- Smoother $\mathbb{P}(v_0 | Y_n) = N(m'_n, C'_n)$
- smoothing is a push forward of filtering under $e^{\mathcal{L} t_n}$

allowing for model error

Note the statistical model used for v may be different from that which generates the data used. We assume that the data we actually incorporate is not y_n but

$$y_n^\epsilon = v_n^\epsilon + \eta_n,$$

where

$$\partial_t v^\epsilon(x,t) + c^\epsilon \cdot \nabla v^\epsilon(x,t) = 0, \quad (x,t) \in \mathbf{T}^2 \times (0,\infty)$$
$$\frac{dv^\epsilon}{dt}(t) + \mathcal{L}^\epsilon v^\epsilon(t) = 0,, \quad t \in (0,\infty).$$

and our filtering/smoothing yields $\mathbb{P}(\cdot|Y_n^\epsilon)$, i.e. Y_n^ϵ replacing Y_n

Questions:

Let $v_0^\epsilon = v_0$ ('true initial condition') and $\delta c = c^\epsilon - c$.

1. large data limit $\lim_{n\to\infty} \mathbb{P}(\cdot|Y_n)$ when $\delta c = 0$;
2. large data limit $\lim_{n\to\infty} \mathbb{P}(\cdot|Y_n^\epsilon =)$ when $0 < |\delta c| \ll 1$?

assumptions

- $\{\phi_k\}$ Fourier basis on $\mathbb{T}^2$.
- $\widehat{C}_0\phi_k = \lambda_k\phi_k$; $\Gamma\phi_k = \gamma_k\phi_k$.
- $\sum_k |k|^{2s}\gamma_k < \infty$; $\gamma_k/\lambda_k = \mathcal{O}(|k|^\beta)$, for some $\beta > 0$.
- $\widehat{m}_0, u \in H^{s+\beta}$.

For $w \in L^2(\mathbb{T}^2)$

$$w = \sum_k w_k\phi_k, \quad w_k = \langle u, \phi_k\rangle$$

then

$$H^\ell = \Big\{w \in L^2(\mathbb{T}^2) \Big| \; \|w\|_\ell^2 := \sum_k |k|^{2\ell}|w_k|^2 < \infty\Big\}.$$

Note that

$$L^2(\mathbb{T}^2) = H^0.$$

perfect model scenario

Theorem

Let $c^\epsilon = c$. Then as $n \to \infty$

$$\mathbb{E}\|\widehat{m}_n - v_n\|_s^2 = \mathcal{O}(n^{-1})$$
$$\mathbb{E}\|m'_n - v_0\|_s^2 = \mathcal{O}(n^{-1})$$
$$\|\widehat{C}_n\|_{\mathcal{L}(L^2,H^s)} = \mathcal{O}(n^{-1}).$$

TRUTH

SMOOTHER

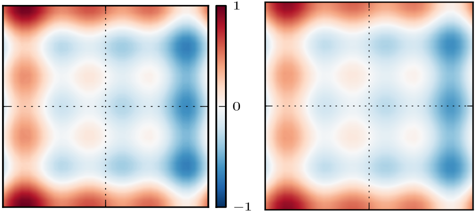


Figure: $v_0 = \sum_{k=1}^{3} \sin(kx) + \cos(ky)$ and $\mathbb{E}(v_0|Y_n)$ for large n

constant velocity difference

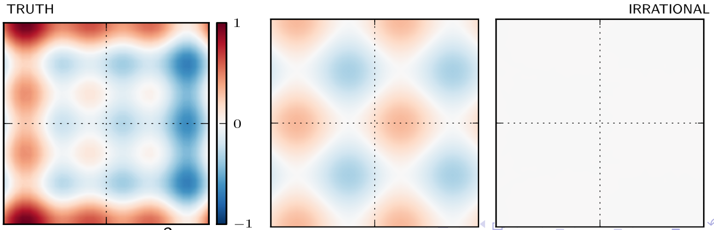
Theorem

Let $c^\epsilon = c + \delta c$. Then as $n \to \infty$

$$\mathbb{E}\|\widehat{m}_n - e^{-t_n\mathcal{L}}\mathcal{M}v_0\|_s^2 = \mathcal{O}(n^{-1})$$
$$\mathbb{E}\|m'_n - \mathcal{M}v_0\|_s^2 = \mathcal{O}(n^{-1})$$

where

$$\mathcal{M} = \sum_{(k_1/p, k_2/q)\in\mathbb{Z}\times\mathbb{Z}}(u,\phi_k)\phi_k \qquad \delta c = (p'/p, q'/q)$$
$$= (u,\phi_0)\ (= \textstyle\int_{\mathbf{T}^2} u\,dxdy = const) \qquad \delta c \in \mathbb{R}\backslash\mathbb{Q}\times\mathbb{R}\backslash\mathbb{Q}$$



Outline
Forward Model
advection equation on a torus
Inverse Problem
data assimilation
allowing for model error
assumptions
Perfect Model Scenario
Model Error
constant velocity difference
integrable velocity difference
white noise velocity difference
Summary

integrable velocity difference

Theorem

Let $c - c^{\epsilon}(t) \to 0$ *and* $\int_0^T (c - c^{\epsilon}(t))\, dt = \alpha + \mathcal{O}(T^{-\kappa})$. *Then as* $n \to \infty$, *for* $\phi = \min\{1, 2\kappa\}$,

$$\mathbb{E}\|\widehat{m}_n - v_n\|_s^2 = \mathcal{O}(n^{-\phi})$$
$$\mathbb{E}\|m'_n - v_{0,\alpha}\|_s^2 = \mathcal{O}(n^{-\phi})$$

where

$$v_{0,\alpha}(\cdot) = v_0(\cdot + \alpha)$$

TRUTH

SMOOTHER

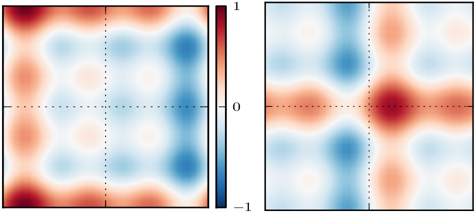


Figure: $v_0 = \sum_{k=1}^{3} \sin(kx) + \cos(ky)$ and $\mathbb{E}(v_0 | Y_n)$ for large n

white noise velocity difference

Theorem

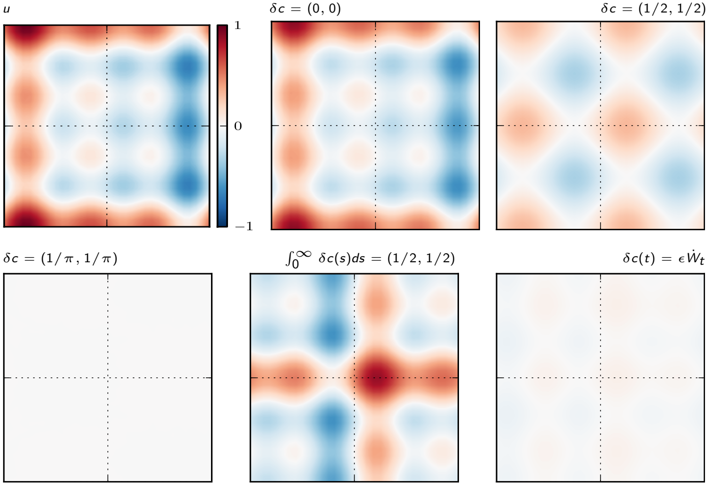
Let $c^\epsilon(t) = c + \epsilon \dot{W}(t)$. *Then as* $n \to \infty$

$$\mathbb{E}\|\widehat{m}_n - \langle u \rangle\|_s^2 = \mathcal{O}(n^{-1})$$
$$\mathbb{E}\|m'_n - \langle u \rangle\|_s^2 = \mathcal{O}(n^{-1})$$

- Here $\langle \cdot \rangle$ denotes the spatial average.
- Similar theorem for different Gaussian perturbations.
- Then obtain a different average.

pictorial overview of theorems

Given $v_0 = \sum_{k=1}^{3} \sin(kx) + \cos(ky)$, $\mathbb{E}(v_0 | Y_n')$ is depicted for large n



verbal overview of results

- We study large data limit of Kalman filter/smoother in infinite dimensions.
- Advection equation on a torus is our forward model.
- Posterior consistency in perfect model scenario.
- Sensitive dependence on wave velocity difference is shown in presence of model error.
- Limits of large data $n \longrightarrow \infty$ and small velocity error $\epsilon \longrightarrow 0$ do not commute.

reference

- "Kalman filtering and smoothing for linear wave equations with model error".
- Wonjung Lee, Damon McDougall and Andrew Stuart.
- To appear: INVERSE PROBLEMS.
- $http://www.maths.warwick.ac.uk/\sim masdr$