Model error in wave propagation problems

Damon McDougall

University of Warwick Mathematics Institute

4th December 2009

Overview

1. Problem set-up (DA for advection)
2. Sampling the posterior distribution
3. Exploring model error in the wave speed
4. A conjecture and acknowledgements

Advection equation

- The system in 2 dimensions:

$$\frac{\partial v}{\partial t} = c \cdot \nabla v, \quad (x,t) \in \mathbb{T}^2 \times [0, T]$$
$$v(x,0) = u(x)$$

- Solution is a continuous translation:

$$v(x,t) = u(x + ct)$$

Advection equation

- The system in 2 dimensions:

$$\frac{\partial v}{\partial t} = c \cdot \nabla v, \quad (x,t) \in \mathbb{T}^2 \times [0,T]$$
$$v(x,0) = u(x)$$

- Solution is a continuous translation:

$$v(x,t) = u(x+ct)$$

1. Torus means periodic boundary conditions on $[0,1] \times [0,1]$
2. Gesture the solution moving in the domain
3. Ask if everybody is happy with the slide

Setting up a Bayesian framework

- Suppose we don't know u exactly but have
  1. A prior on u
  $$\mathbb{P}(u) = \mathcal{N}(0, (-\Delta)^{-\alpha})$$
  2. Observations $y_{j,k} = v(x_j, t_k) + \eta_{j,k},\ \eta_{j,k} \sim \mathcal{N}(0, \gamma^2)$, i.e.,
  $$y = G(u) + \eta,\ \eta \sim \mathcal{N}(0, \Gamma)$$
- Aim is to get information from the posterior:
$$\mathbb{P}(u|y) \propto \mathbb{P}(y|u)\mathbb{P}(u)$$

Setting up a Bayesian framework

- Suppose we don't know u exactly but have
  1. A prior on u
  $$\mathbb{P}(u) = \mathcal{N}(0, (-\Delta)^{-\alpha})$$
  2. Observations $y_{j,k} = v(x_j, t_k) + \eta_{j,k}, \ \eta_{j,k} \sim \mathcal{N}(0, \gamma^2)$, i.e.,
  $$y = G(u) + \eta, \ \eta \sim \mathcal{N}(0, \Gamma)$$
- Aim is to get information from the posterior:
$$\mathbb{P}(u|y) \propto \mathbb{P}(y|u)\mathbb{P}(u)$$

1. If we know u and the model then we know everything in time. So suppose we don't know u.
2. Covariance operator explanation. Compare with finite dimensional analogue (pds matrix vs discretised Laplacian). Mention that covariance in infinite dimensions gives smoothness of function
3. Can write $v(x_j, t_k)$ as a (in this case) linear function of the initial condition u
4. This is just Bayes' theorem
5. I'm now going to show you how to sample the posterior
6. Ask if everybody is happy with the slide

Sampling $\mathbb{P}(u|y)$ (MCMC)

- A draw, ξ, from the prior, $\mathcal{N}(0, (-\Delta)^{-\alpha})$, looks like

$$\xi = \sum_{k \in \mathbb{Z}^2 \setminus \{0\}} \frac{\gamma_k}{(4\pi^2 |k|^2)^{\alpha/2}} \exp(2\pi i k \cdot x)$$

- Construct a Markov chain $\{u_j\}_{j=1}^{\infty}$.
  1. Let u_n be current state and make a proposal w:
  $$w = (1 - \beta^2)^{1/2} u_n + \beta \xi, \quad \text{some } \beta \in (0, 1)$$
  2. Define $\Phi(\cdot) := \frac{1}{2} \left|\left| G(\cdot) - y \right|\right|_{\Gamma}^2$ and set $u_{n+1} = w$ with probability
  $$\alpha(u_n, w) = \min \left\{1, \exp\left(\Phi(u_n) - \Phi(w)\right)\right\}$$
  and $u_{n+1} = u_n$ otherwise.

Sampling $\mathbb{P}(u|y)$ (MCMC)

- A draw, ξ, from the prior, $\mathcal{N}(0,(-\Delta)^{-\alpha})$, looks like

$$\xi = \sum_{k \in \mathbb{Z}^2 \setminus \{0\}} \frac{\gamma_k}{(4\pi^2|k|^2)^{\alpha/2}} \exp(2\pi i k \cdot x)$$

- Construct a Markov chain $\{u_j\}_{j=1}^{\infty}$.
  1. Let u_n be current state and make a proposal w:
  $$w = (1-\beta^2)^{1/2} u_n + \beta\xi, \quad \text{some } \beta \in (0,1)$$
  2. Define $\Phi(\cdot) := \frac{1}{2}||G(\cdot) - y||_\Gamma^2$ and set $u_{n+1} = w$ with probability
  $$\alpha(u_n, w) = \min\{1, \exp(\Phi(u_n) - \Phi(w))\}$$
  and $u_{n+1} = u_n$ otherwise.

Tell people not to panic about this slide

1. To sample from the posterior distribution: Need to sample from the prior and set up a Markov chain with the posterior measure as the invariant measure
2. Link the decay of Fourier coefficients (and smoothness) to the covariance operator in infinite dimensions
3. Finding a sensible β. Effect of small/large β
4. Give intuition: Explain what happens when proposed state matches the data better/worse than the current state
5. So chain is attracted to modes - natural behaviour for sampling modes. It is unlikely to explore the tails.
6. Ask if everybody is happy with the slide
7. Show movie 256-burnin.avi

Model error set-up

- We explore what happens when the data y is observed from the model

$$\frac{\partial v}{\partial t} = c \cdot \nabla v$$

as before, but we don't know what c is.

- We use a perturbed model for what we *believe* y came from

$$\frac{\partial v}{\partial t} = c' \cdot \nabla v$$

- What is the effect of a mismatch between the true c and the c' used for assimilation?

Model error set-up

- We explore what happens when the data y is observed from the model

$$\frac{\partial v}{\partial t} = c \cdot \nabla v$$

as before, but we don't know what c is.

- We use a perturbed model for what we *believe* y came from

$$\frac{\partial v}{\partial t} = c' \cdot \nabla v$$

- What is the effect of a mismatch between the true c and the c' used for assimilation?

1. Now explore what happens when we don't know the model exactly
2. c is 'true' wave speed
3. c' is what we use in our assimilation (it is incorrect)
4. Scenario: Someone gives us y, we know nothing about c. So we take a guess, c'
5. We explore what will happen when we vary c'
6. Ask if everybody is happy with the slide

Wave speed interpolation

- The approach taken was to linearly interpolate between two wave speeds

$$(0.5, 1.0) = c, c_1, c_2, \cdots, c_8, c_9, c' = (0.475, 0.95)$$

- We looked at

$$\bar{u}_N = \left\| \frac{1}{N} \sum_{n=1}^{N} u_n \right\|$$

for N rather large ($\approx 10 \times 10^6$) as you vary through the linear interpolation above.

Wave speed interpolation

- The approach taken was to linearly interpolate between two wave speeds

$$(0.5, 1.0) = c, c_1, c_2, \cdots, c_8, c_9, c' = (0.475, 0.95)$$

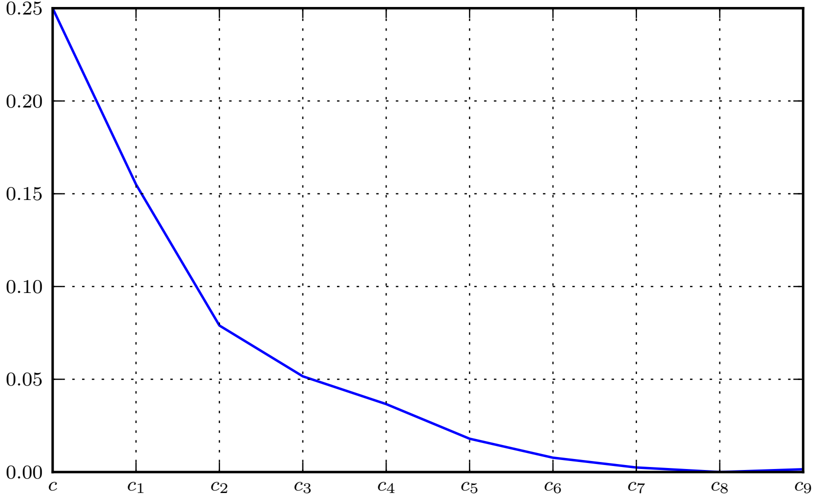
- We looked at

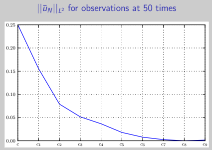
$$\bar{u}_N = \left\| \frac{1}{N} \sum_{n=1}^{N} u_n \right\|$$

for N rather large ($\approx 10 \times 10^6$) as you vary through the linear interpolation above.

1. This is the empircal mean of samples from the chain
2. Ask if everybody is happy with this slide

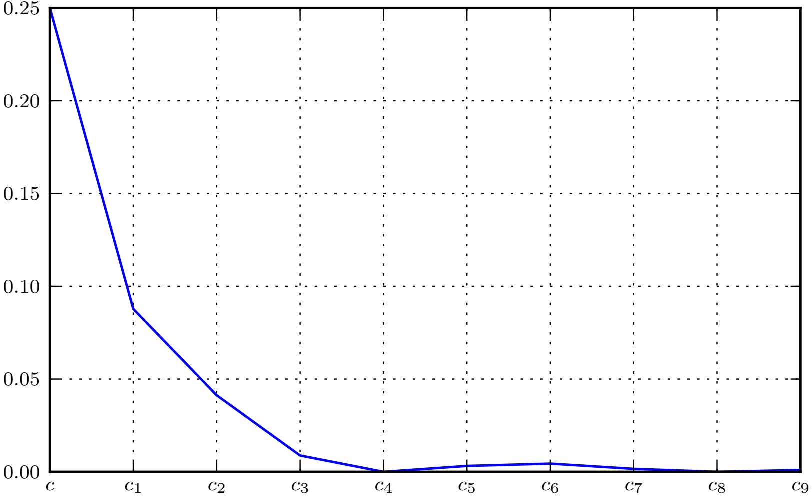
$||\bar{u}_N||_{L^2}$ for observations at 50 times



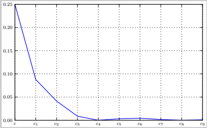


1. Observations taken at 256 points in space, 50 in time
2. Blue line is L^2 norm of $\bar{u}_N$
3. Show graphs
4. Does everybody understand the graph?

$||\bar{u}_N||_{L^2}$ for observations at 100 times

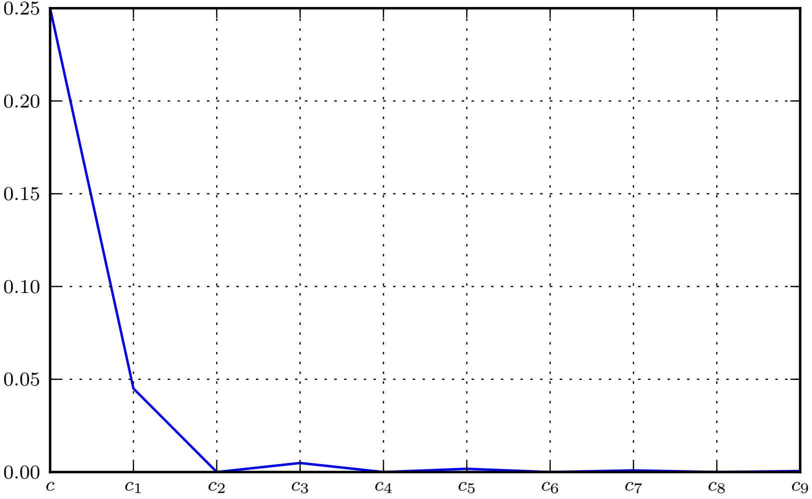


$||\bar{u}_N||_{L^2}$ for observations at 100 times

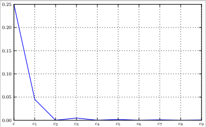


1. Observations taken at 256 points in space, 100 in time
2. Show graphs
3. Blue line is L^2 norm of $\bar{u}_N$

$||\bar{u}_N||_{L^2}$ for observations at 200 times

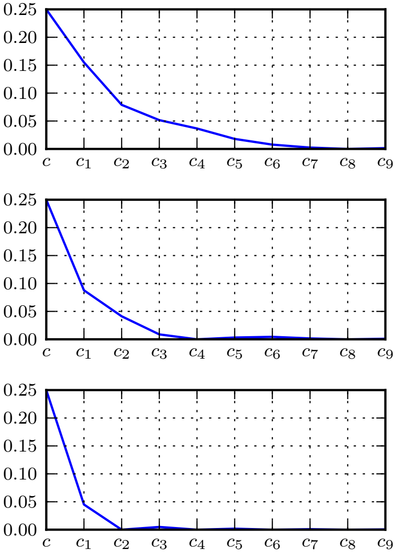


$||\bar{u}_N||_{L^2}$ for observations at 200 times

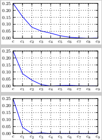


1. Observations taken at 256 points in space, 200 in time
2. Say no graphs, but they're pretty much the same
3. Blue line is L^2 norm of $\bar{u}_N$

Comparing $||\bar{u}_N||_{L^2}$ for 50, 100 and 200 times



Comparing $||\bar{u}_N||_{L^2}$ for 50, 100 and 200 times



1. Comparison of all three

A conjecture

- Suppose we make observations at K times. Define $\bar{u}$ to be the mean of $\mathbb{P}(u|y)$, then

$$\forall c' \neq c, \; ||\bar{u}||_{L^2} \to 0 \text{ as } K \to \infty$$

- This result is being investigated by Wonjung Lee (NERC funded)
- Acknowledgements go to:
    1. NERC for funding
    2. Warwick CSC for computing time
    3. Wonjung Lee for analysis
    4. Andrew Stuart and Chris Jones for patience, help and encouragement
- Questions?

A conjecture

- Suppose we make observations at K times. Define $\bar{u}$ to be the mean of $\mathbb{P}(u|y)$, then

$$\forall c' \neq c, \ ||\bar{u}||_{L^2} \to 0 \text{ as } K \to \infty$$

- This result is being investigated by Wonjung Lee (NERC funded)
- Acknowledgements go to:
  1. NERC for funding
  2. Warwick CSC for computing time
  3. Wonjung Lee for analysis
  4. Andrew Stuart and Chris Jones for patience, help and encouragement
- Questions?

1. Now make a conjecture based on the graphs
2. This is saying that if you get the wave speed wrong, then L^2 norm of the posterior mean goes to zero if you wait long enough
3. My PhD position, Wonjung Lee's post-doctoral position
4. All computations done on a ≈ 900 processor supercomputer hidden in a secret location called Physical Sciences