Bayesian data assimilation for a toy model

Damon McDougall

Mathematics Institute, University of Warwick, United Kingdom

9th June 2010

Overview

1. Problem set-up
2. The ‘prior’ distribution
3. Optimiser’s and sampler’s approach
4. Model error
5. ‘Lagrangian’ observations and links to oceanography
6. Cool video

The problem

- Given

$$
\begin{aligned}
\text{(PDE)} \quad & \frac{\partial v}{\partial t} = c \cdot \nabla v, \quad (x,t) \in \mathbb{T}^2 \times [0, T], \text{ and} \\
\text{(IC)} \quad & v(x,0) = u(x)
\end{aligned}
$$

The problem

- Given

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

find $v : \mathbb{T}^2 \times [0,T] \to \mathbb{R}$ satisfying (PDE) and (IC)

- Note: the solution to (PDE) is $v(x,t) = u(x+ct)$

The problem

- Given

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

find $v : \mathbb{T}^2 \times [0,T] \to \mathbb{R}$ satisfying (PDE) and (IC)

$u : \mathbb{T}^2 \to \mathbb{R}$ is called the *initial condition*

- Note: the solution to (PDE) is $v(x,t) = u(x+ct)$

The problem

- Given

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

find $v : \mathbb{T}^2 \times [0, T] \to \mathbb{R}$ satisfying (PDE) and (IC)

$u : \mathbb{T}^2 \to \mathbb{R}$ is called the *initial condition*

- This is called a *forward* problem

- Note: the solution to (PDE) is $v(x,t) = u(x + ct)$

The problem

- Given

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

$$\text{(IC)} \quad \cancel{v(x, 0) = u(x)}$$

find $v : \mathbb{T}^2 \times [0, T] \to \mathbb{R}$ satisfying (PDE) and (IC)

$u : \mathbb{T}^2 \to \mathbb{R}$ is called the *initial condition*

- Note: the solution to (PDE) is $v(x, t) = u(x + ct)$

The problem

- Given

$$
\begin{aligned}
\text{(PDE)} \quad & \frac{\partial v}{\partial t} = c \cdot \nabla v, \quad (x,t) \in \mathbb{T}^2 \times [0, T], \text{ and} \\
\text{(IC)} \quad & \cancel{v(x,0) = u(x)}
\end{aligned}
$$

and observations

$$
y_{j,k} = v(x_j, t_k) + \eta_{j,k}, \quad \eta_{j,k} \overset{\text{i.i.d}}{\sim} \mathcal{N}(0, \sigma^2)
$$

$u : \mathbb{T}^2 \to \mathbb{R}$ is called the *initial condition*

The problem

- Given

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

and observations

$$y_{j,k} = v(x_j, t_k) + \eta_{j,k}, \quad \eta_{j,k} \overset{\text{i.i.d}}{\sim} \mathcal{N}(0,\sigma^2)$$

Find $u : \mathbb{T}^2 \to \mathbb{R}$

The problem

- Given

$$\begin{aligned}
&\text{(PDE)} \quad \frac{\partial v}{\partial t} = c \cdot \nabla v, \quad (x,t) \in \mathbb{T}^2 \times [0,T], \text{ and} \\
&\text{(IC)} \quad \cancel{v(x,0) = u(x)}
\end{aligned}$$

and observations

$$\begin{aligned}
& y_{j,k} = v(x_j, t_k) + \eta_{j,k}, \quad \eta_{j,k} \overset{\text{i.i.d}}{\sim} \mathcal{N}(0, \sigma^2) \\
\rightsquigarrow \quad & y = \mathcal{G}(u) + \eta, \quad \eta \sim \mathcal{N}(0, B)
\end{aligned}$$

Find $u : \mathbb{T}^2 \to \mathbb{R}$

The problem

- Given

$$\begin{aligned} \text{(PDE)} \quad & \frac{\partial v}{\partial t} = c \cdot \nabla v, \quad (x,t) \in \mathbb{T}^2 \times [0,T], \text{ and} \\ \text{(IC)} \quad & \cancel{v(x,0) = u(x)} \end{aligned}$$

  and observations

$$\begin{aligned} & y_{j,k} = v(x_j, t_k) + \eta_{j,k}, \quad \eta_{j,k} \overset{\text{i.i.d}}{\sim} \mathcal{N}(0, \sigma^2) \\ \rightsquigarrow \quad & y = \mathcal{G}(u) + \eta, \quad \eta \sim \mathcal{N}(0, B) \end{aligned}$$

  Find $u : \mathbb{T}^2 \to \mathbb{R}$

- This is called an *inverse* problem

The problem

- Given

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

and observations

$$y_{j,k} = v(x_j, t_k) + \eta_{j,k}, \quad \eta_{j,k} \overset{\text{i.i.d}}{\sim} \mathcal{N}(0,\sigma^2)$$
$$\rightsquigarrow \quad y = \mathcal{G}(u) + \eta, \quad \eta \sim \mathcal{N}(0,B)$$

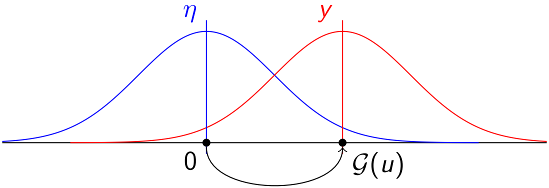
Find $u : \mathbb{T}^2 \to \mathbb{R}$

- This is called an *inverse* problem
- Data Assimilation is act of incorporating y into (PDE) to get u

The posterior

- Given y, find u. I.e., want to know what $\mathbb{P}(u|y)$ 'looks like'
- $\mathbb{P}(u|y)$ is called the *posterior* ('after' the data) distribution

$$\begin{aligned}\mathbb{P}(u|y) &= \frac{\mathbb{P}(y|u)\mathbb{P}(u)}{\mathbb{P}(y)} \\ &\propto \mathbb{P}(y|u)\mathbb{P}(u)\end{aligned}$$



- $\mathcal{G}$ must be linear (see why later)

Gaussians

- Fact: Gaussian probability distribution function (pdf) has the form

$$\exp(-(ax^2 + bx + c))$$

- Note: Product of two Gaussians is Gaussian

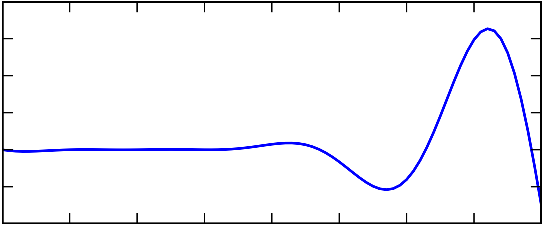
$$\begin{aligned} &\exp(-(a_1x^2 + b_1x + c_1))\exp(-(a_2x^2 + b_2x + c_2)) \\ = &\exp(-((a_1 + a_2)x^2 + (b_1 + b_2)x + (c_1 + c_2))) \end{aligned}$$

- So

$$\mathbb{P}(y|u), \mathbb{P}(u) \text{ Gaussian} \Rightarrow \mathbb{P}(y|u)\mathbb{P}(u) \propto \mathbb{P}(u|y) \text{ Gaussian}$$

The prior

- We need $\mathbb{P}(u)$. Called the *prior* ('before' the data) distribution
- 'Prior knowledge' means have idea of some property of u. E.g., Temperature vs. Time:



Might expect this to have, say, one derivative.
This 'prior knowledge' is usually either

1. Given to us
2. Obtained from past experience
3. A complete guess

The prior

- How do we construct a 'random function' with, say, one derivative?
-

The prior

- How do we construct a ‘random function’ with, say, one derivative?
-

$$\xi(x) = \sum_{k \in \mathbb{K}^2} \frac{\qquad\qquad}{\qquad\qquad} \exp(2\pi i k \cdot x),$$

The prior

- How do we construct a 'random function' with, say, one derivative?
- Let $\mathbb{K}^2 = \mathbb{Z}^2 \setminus \{(0,0)\}$

$$\xi(x) = \sum_{k \in \mathbb{K}^2} \frac{}{(4\pi^2 |k|^2)} \exp(2\pi i k \cdot x),$$

The prior

- How do we construct a 'random function' with, say, one derivative?
- Let $\mathbb{K}^2 = \mathbb{Z}^2 \setminus \{(0,0)\}$

$$\xi(x) = \sum_{k \in \mathbb{K}^2} \frac{\gamma_k}{(4\pi^2 |k|^2)} \exp(2\pi i k \cdot x),$$

The prior

- How do we construct a 'random function' with, say, one derivative?
- Let $\mathbb{K}^2 = \mathbb{Z}^2 \setminus \{(0,0)\}$, a random function looks like

$$\xi(x) = \sum_{k \in \mathbb{K}^2} \frac{\gamma_k}{(4\pi^2 |k|^2)} \exp(2\pi i k \cdot x), \quad \gamma_k \overset{\text{i.i.d}}{\sim} \mathcal{N}(0, I_2)$$

The prior

- How do we construct a 'random function' with, say, $\alpha - 1$ derivatives?
- Let $\mathbb{K}^2 = \mathbb{Z}^2 \setminus \{(0,0)\}$, a random function looks like

$$\xi(x) = \sum_{k \in \mathbb{K}^2} \frac{\gamma_k}{(4\pi^2|k|^2)^{\alpha/2}} \exp(2\pi i k \cdot x), \quad \gamma_k \overset{\text{i.i.d}}{\sim} \mathcal{N}(0, I_2)$$

- $\xi \in H^{\alpha-1}$

The prior

- How do we construct a 'random function' with, say, $\alpha - 1$ derivatives?
- Let $\mathbb{K}^2 = \mathbb{Z}^2 \setminus \{(0,0)\}$, a random function looks like

$$\xi(x) = \sum_{k \in \mathbb{K}^2} \frac{\gamma_k}{(4\pi^2 |k|^2)^{\alpha/2}} \exp(2\pi i k \cdot x), \quad \gamma_k \overset{\text{i.i.d}}{\sim} \mathcal{N}(0, I_2)$$

- $\xi \in H^{\alpha-1}$
- Notation for the distribution of functions drawn in this manner is

$$\mathcal{N}(0, (-\Delta)^{-\alpha}), \qquad \text{where } -\Delta = -\frac{\mathrm{d}^2}{\mathrm{d}x^2}$$

The prior

- How do we construct a 'random function' with, say, $\alpha - 1$ derivatives?
- Let $\mathbb{K}^2 = \mathbb{Z}^2 \setminus \{(0,0)\}$, a random function looks like

$$\xi(x) = \sum_{k \in \mathbb{K}^2} \frac{\gamma_k}{(4\pi^2|k|^2)^{\alpha/2}} \exp(2\pi i k \cdot x), \quad \gamma_k \overset{\text{i.i.d}}{\sim} \mathcal{N}(0, I_2)$$

- $\xi \in H^{\alpha-1}$
- Notation for the distribution of functions drawn in this manner is

$$\mathcal{N}(0, (-\Delta)^{-\alpha}), \qquad \text{where } -\Delta = -\frac{\mathrm{d}^2}{\mathrm{d}x^2}$$

- The prior $\mathbb{P}(u)$ will be exactly this. It is Gaussian

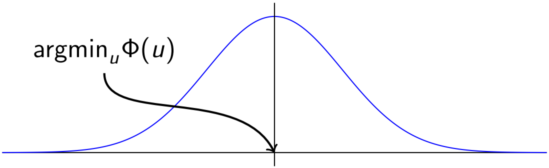
Finding 'the answer'

- $f : \mathcal{S} \to \mathbb{R}$ a probability density function if

$$f(x) \geq 0 \quad \forall x \in \mathcal{S} \qquad \text{and} \qquad \int_{\mathcal{S}} f(x)\,\mathrm{d}x = 1$$

- Define $\Phi(\cdot) := \frac{1}{2}\left\|\mathcal{G}(\cdot) - y\right\|_B^2$. Objective is to minimise this
- Can turn into a probability density function

$$\exp(-\Phi(\cdot)) = \exp\left(-\frac{1}{2}\left\|\mathcal{G}(\cdot) - y\right\|_B^2\right)$$



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

- Idea: Construct $\{u_j\}_{j=1}^{\infty}$ cleverly such that $\{u_j\}_{j=1}^{\infty} \overset{\text{i.i.d}}{\sim} \mathbb{P}(u|y)$
    1. Let u_k be the 'current' state in the sequence and construct a *proposal*, w

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

- Idea: Construct $\{u_j\}_{j=1}^{\infty}$ cleverly such that $\{u_j\}_{j=1}^{\infty} \overset{\text{i.i.d}}{\sim} \mathbb{P}(u|y)$
    1. Let u_k be the 'current' state in the sequence and construct a *proposal*, w
    2. Define $\Phi(\cdot) := \frac{1}{2} \left\| \mathcal{G}(\cdot) - y \right\|_B^2$

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

- Idea: Construct $\{u_j\}_{j=1}^{\infty}$ cleverly such that $\{u_j\}_{j=1}^{\infty} \overset{\text{i.i.d}}{\sim} \mathbb{P}(u|y)$
  1. Let u_k be the 'current' state in the sequence and construct a *proposal*, w
  2. Define $\Phi(\cdot) := \frac{1}{2} \|\mathcal{G}(\cdot) - y\|_B^2$
  3. Compute $\alpha(u_k, w) = 1 \wedge \exp(\Phi(u_k) - \Phi(w))$

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

- Idea: Construct $\{u_j\}_{j=1}^{\infty}$ cleverly such that $\{u_j\}_{j=1}^{\infty} \overset{\text{i.i.d}}{\sim} \mathbb{P}(u|y)$
  1. Let u_k be the 'current' state in the sequence and construct a *proposal*, w
  2. Define $\Phi(\cdot) := \frac{1}{2} \|\mathcal{G}(\cdot) - y\|_B^2$
  3. Compute $\alpha(u_k, w) = 1 \wedge \exp(\Phi(u_k) - \Phi(w))$
  4. Let

$$u_{k+1} = \begin{cases} w & \text{with probability } \alpha(u_k, w) \\ u_k & \text{with probability } 1 - \alpha(u_k, w) \end{cases}$$

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

- Idea: Construct $\{u_j\}_{j=1}^{\infty}$ cleverly such that $\{u_j\}_{j=1}^{\infty} \overset{\text{i.i.d}}{\sim} \mathbb{P}(u|y)$

 1. Let u_k be the 'current' state in the sequence. Make a draw $\xi \sim \mathcal{N}(0, (-\Delta)^{-\alpha})$ and construct a *proposal*, w

 $$w = (1-\beta^2)^{\frac{1}{2}} u_k + \beta\xi, \quad \text{some } \beta \in (0,1)$$

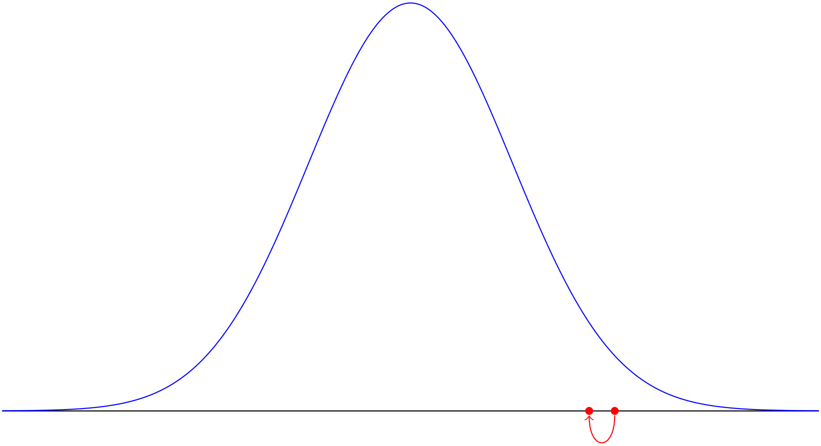
 2. Define $\Phi(\cdot) := \frac{1}{2} \|\mathcal{G}(\cdot) - y\|_B^2$
 3. Compute $\alpha(u_k, w) = 1 \wedge \exp(\Phi(u_k) - \Phi(w))$
 4. Let

 $$u_{k+1} = \begin{cases} w & \text{with probability } \alpha(u_k, w) \\ u_k & \text{with probability } 1 - \alpha(u_k, w) \end{cases}$$

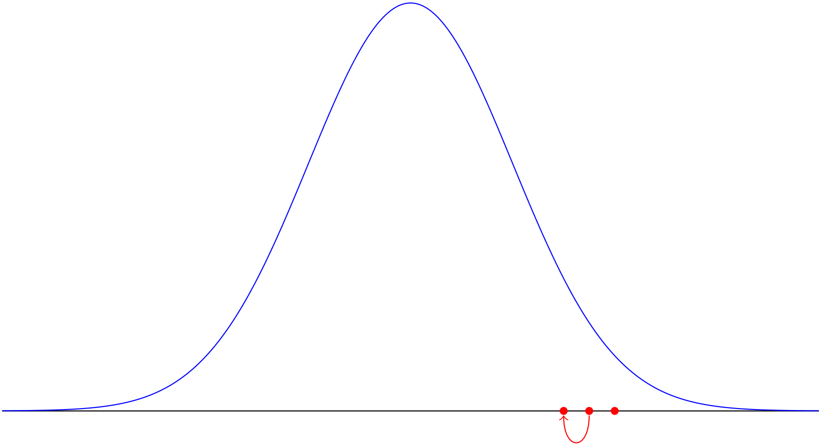
Sampling $\mathbb{P}(u|y)$

- Idea: Construct $\{u_j\}_{j=1}^{\infty}$ cleverly such that $\{u_j\}_{j=1}^{\infty} \overset{\text{i.i.d}}{\sim} \mathbb{P}(u|y)$
  1. Let u_k be the 'current' state in the sequence. Make a draw $\xi \sim \mathcal{N}(0, (-\Delta)^{-\alpha})$ and construct a *proposal*, w
  $$w = (1 - \beta^2)^{\frac{1}{2}} u_k + \beta\xi, \quad \text{some } \beta \in (0, 1)$$
  2. Define $\Phi(\cdot) := \frac{1}{2} \|\mathcal{G}(\cdot) - y\|_B^2$
  3. Compute $\alpha(u_k, w) = 1 \wedge \exp(\Phi(u_k) - \Phi(w))$
  4. Let
  $$u_{k+1} = \begin{cases} w & \text{with probability } \alpha(u_k, w) \\ u_k & \text{with probability } 1 - \alpha(u_k, w) \end{cases}$$
- Take u_1 to be a draw from $\mathcal{N}(0, (-\Delta)^{-\alpha})$

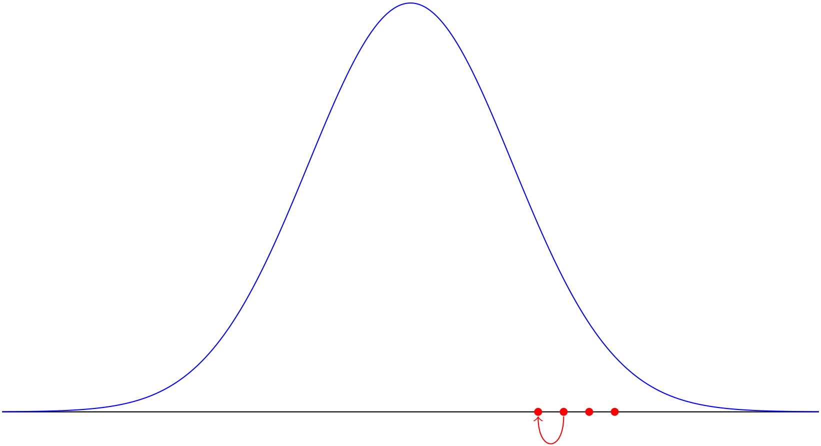
Sampling $\mathbb{P}(u|y)$



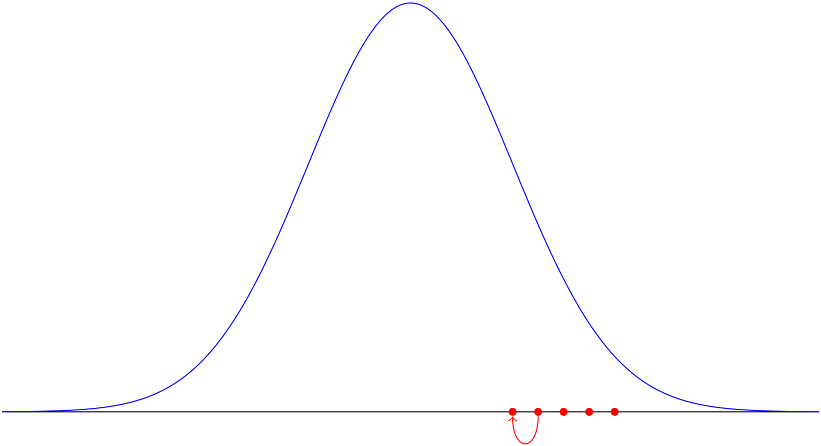
Sampling $\mathbb{P}(u|y)$



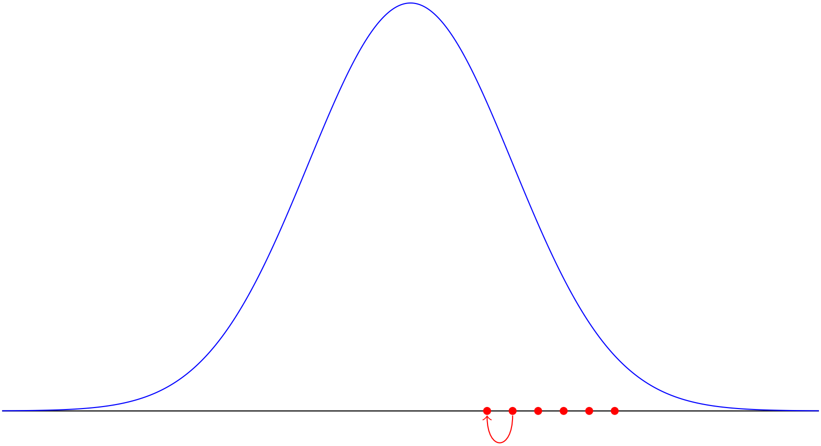
Sampling $\mathbb{P}(u|y)$



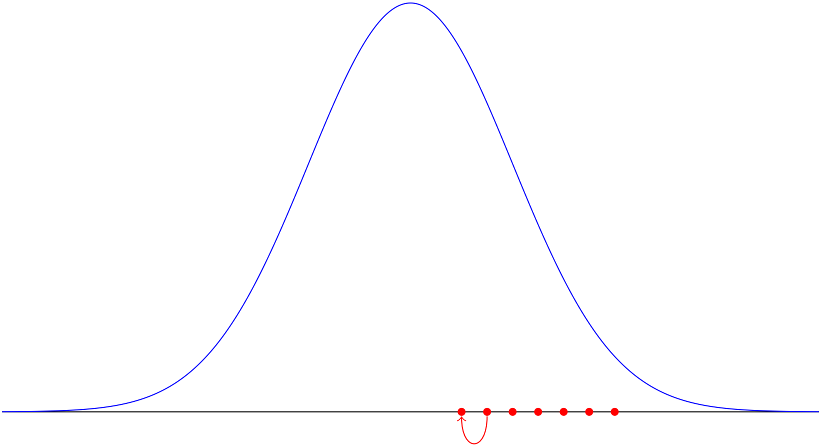
Sampling $\mathbb{P}(u|y)$



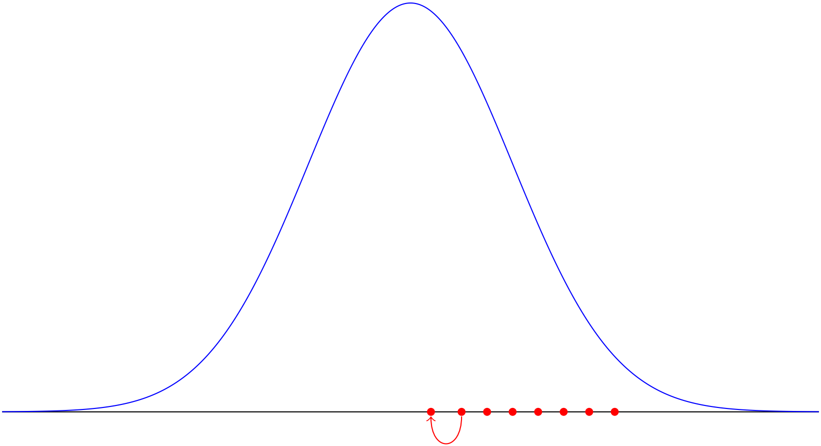
Sampling $\mathbb{P}(u|y)$



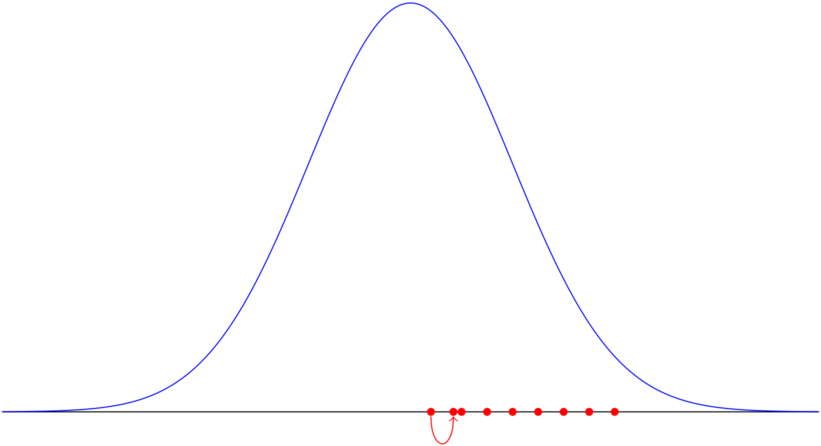
Sampling $\mathbb{P}(u|y)$



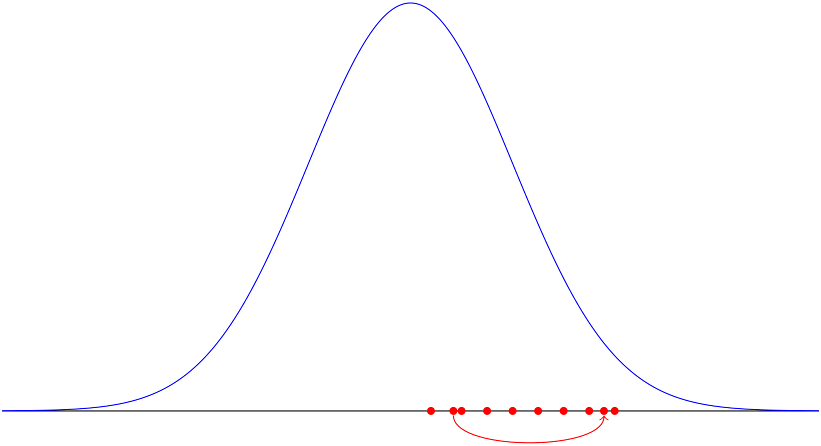
Sampling $\mathbb{P}(u|y)$



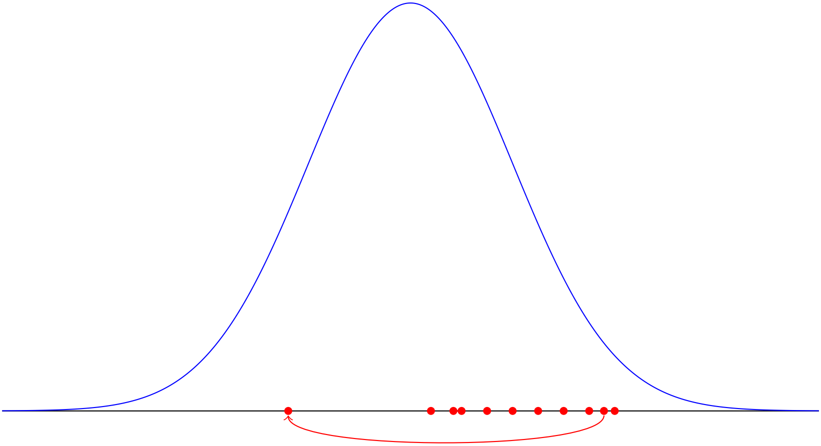
Sampling $\mathbb{P}(u|y)$



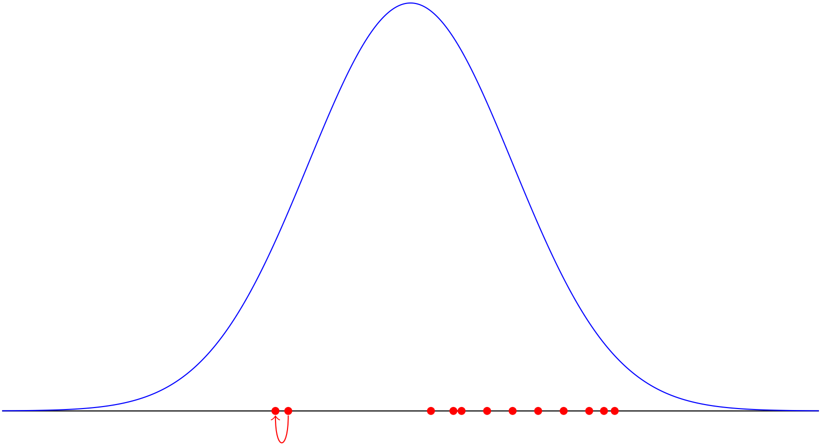
Sampling $\mathbb{P}(u|y)$



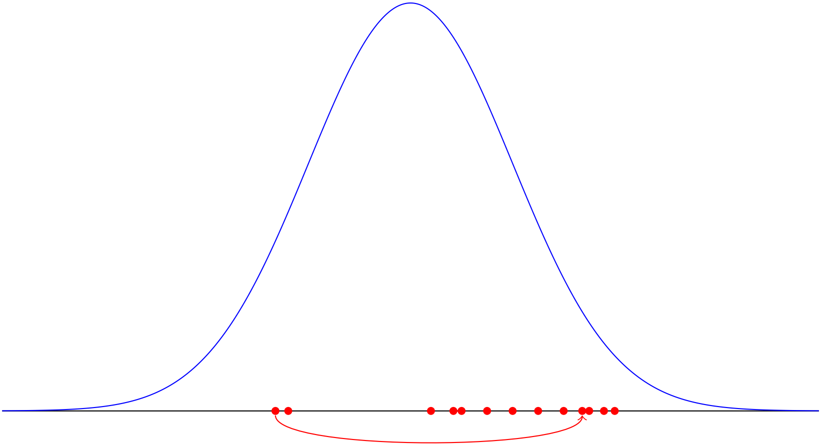
Sampling $\mathbb{P}(u|y)$



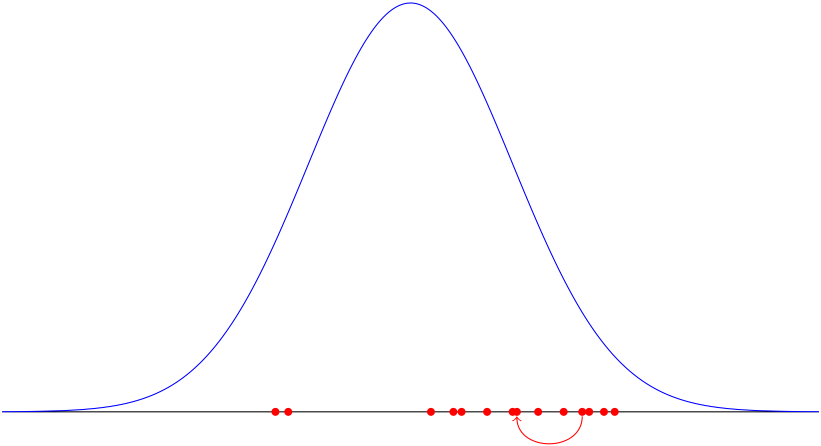
Sampling $\mathbb{P}(u|y)$



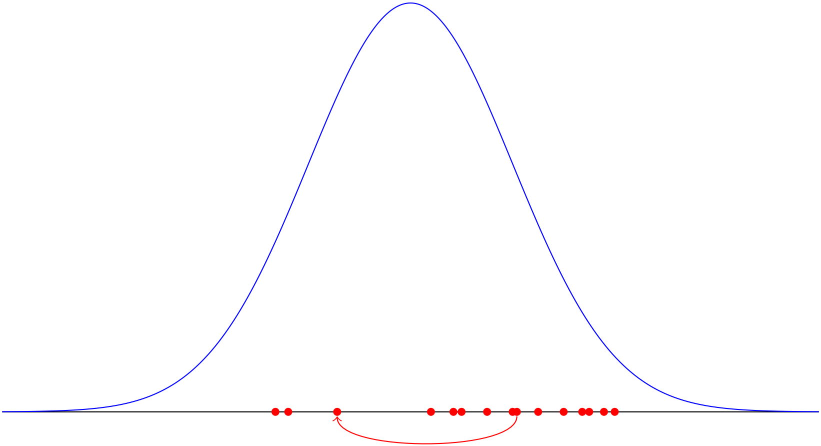
Sampling $\mathbb{P}(u|y)$



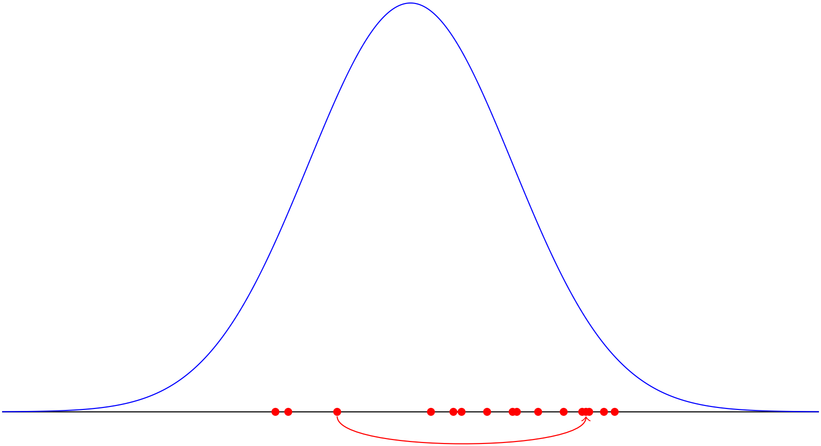
Sampling $\mathbb{P}(u|y)$



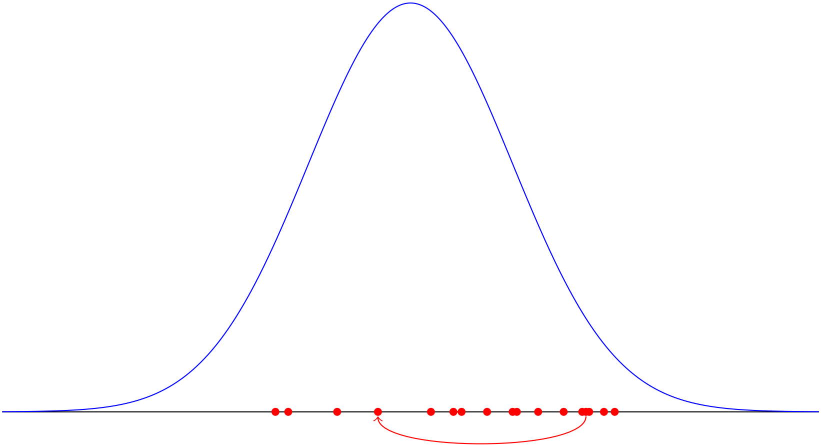
Sampling $\mathbb{P}(u|y)$



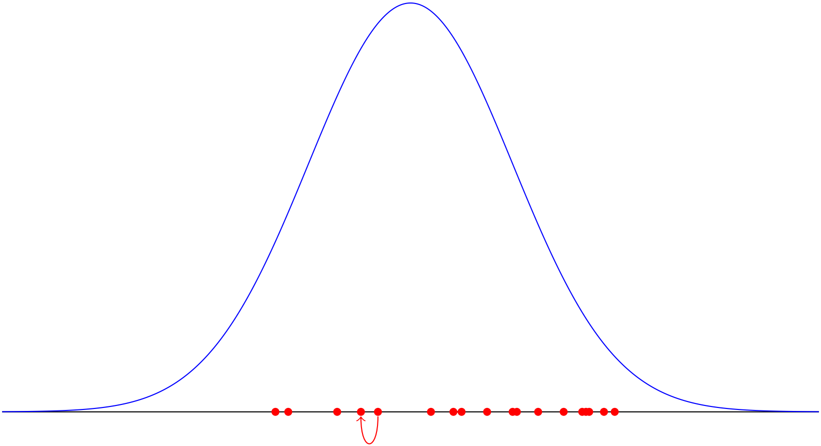
Sampling $\mathbb{P}(u|y)$



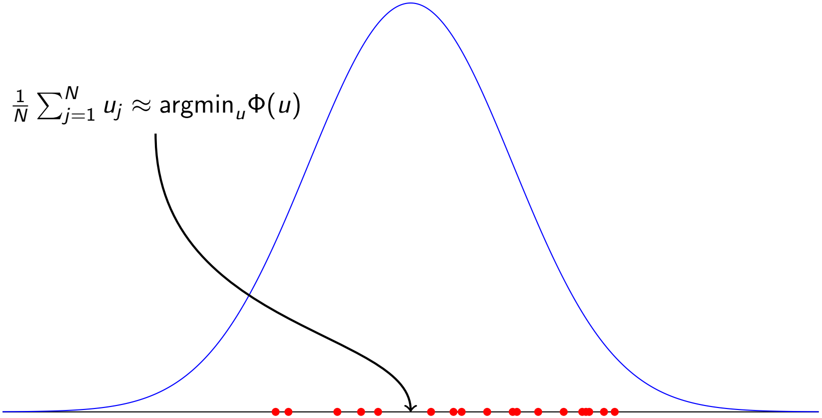
Sampling $\mathbb{P}(u|y)$



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



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



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

- Back to our model

$$\begin{aligned} \text{(PDE)} \quad & \frac{\partial v}{\partial t} = c \cdot \nabla v, \quad (\underline{x}, t) \in \mathbb{T}^2 \times [0, T], \text{ and} \\ \text{(IC)} \quad & v(\underline{x}, 0) = u(\underline{x}) \end{aligned}$$

- The initial condition will be

$$u(\underline{x}) = \sin(2\pi x)\cos(2\pi y)$$

- Make observations $y_{jk} = v(x_j, t_k) + \eta_{jk}$
- Now pretend we don't know u
- See if we can recover u from y by sampling $\mathbb{P}(u|y)$

Model error

- What happens if y does not come from the model you use?
- There are really two models:
  1. The *true model*: One from which observations are made
  2. The *process model*: One which is used to compute $\Phi(\cdot) := \frac{1}{2} \|\mathcal{G}(\cdot) - y\|_B^2$
- This happens in climate science

Model error

- Let's say we have no idea what the wavespeed, c, is

- The **true** model:

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

- The **process** model:

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

- We are given $y_{jk} = v(x_j, t_k) + \eta_{jk}$, with v from the **true** model

1. Sample $\{u_j\}_{j=1}^n \sim \mathbb{P}(u|y)$. Use **process** model to evaluate Φ

Model error

- Let's say we have no idea what the wavespeed, c, is

- The **true** model:

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

- The **process** model:

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

- We are given $y_{jk} = v(x_j, t_k) + \eta_{jk}$, with v from the **true** model

1. Sample $\{u_j\}_{j=1}^n \sim \mathbb{P}(u|y)$. Use **process** model to evaluate Φ
2. Look at $\sum_{j=1}^n u_j$ for some c.

Model error

- Let's say we have no idea what the wavespeed, c, is

- The **true** model:

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

- The **process** model:

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

- We are given $y_{jk} = v(x_j, t_k) + \eta_{jk}$, with v from the **true** model

1. Sample $\{u_j\}_{j=1}^n \sim \mathbb{P}(u|y)$. Use **process** model to evaluate Φ
2. Look at $\sum_{j=1}^n u_j$ for some c.
3. Change c then go to 1.

Model error

- Let's say we have no idea what the wavespeed, c, is

- The **true** model:

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

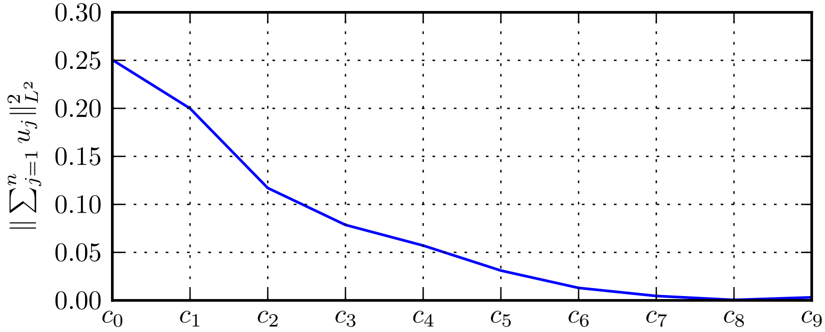
- The **process** model:

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

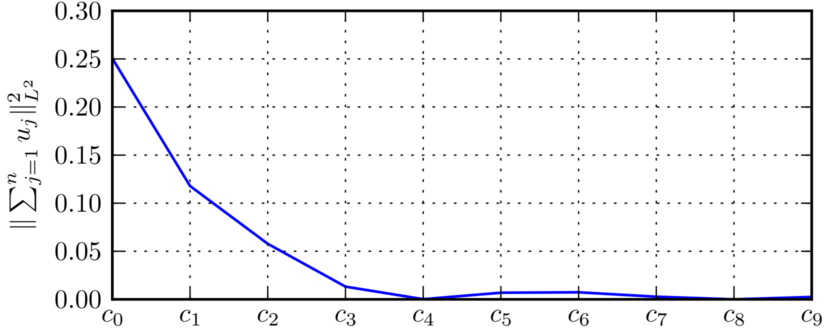
- We are given $y_{jk} = v(x_j, t_k) + \eta_{jk}$, with v from the **true** model

1. Sample $\{u_j\}_{j=1}^n \sim \mathbb{P}(u|y)$. Use **process** model to evaluate Φ
2. Look at $\left\|\sum_{j=1}^n u_j\right\|_{L^2}^2$ for some c.
3. Change c then go to 1.

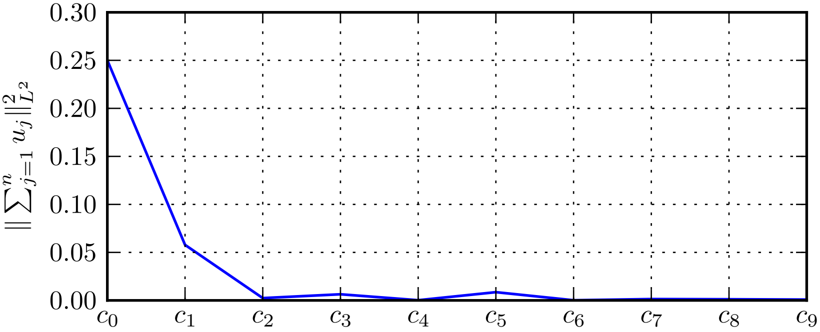
Model error: 50 observations



Model error: 100 observations



Model error: 200 observations



Model error

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

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

Model error

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

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

- This is not exactly correct. Why?

Model error

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

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

- This is not exactly correct. Why?
- Correct up to time periodicity. Also called 'aliasing'

Lagrangian Data Assimilation

- 'Eulerian' obervations:

$$y_{jk} = v(x_j, t_k) + \eta_{jk}, \quad \eta_{jk} \overset{\text{i.i.d}}{\sim} \mathcal{N}(0, \sigma^2)$$

- Set

$$\frac{\mathrm{d}z}{\mathrm{d}t} = v(z, t)$$

- 'Lagrangian' observations:

$$y_k = z(t_k) + \eta_k, \quad \eta_k \overset{\text{i.i.d}}{\sim} \mathcal{N}(0, \Gamma)$$

- Qualitatively, $z(t) \in \Omega \subseteq \mathbb{R}^2$ are positions of massless rubber ducks
- These observations are more commonly used in oceanography

Summary

- Bayes' Rule applied to Inverse Problems
- Optimiser's and sampler's approach to Data Assimilation
- Random functions and the prior
- Model error and the problems it causes
- Eulerian and Lagrangian observations
- Cool video

Thank you