Bayesian inference for inverse obstacle scattering problems

Kate Lee, Colin Fox, and Geoff Nicholls

School of Computing and Mathematical Science, AUT
Department of Physics, University of Otago
Department of Statistics, University of Oxford

Southern Uncertainty Quantification, January 7-9, 2013,
University of Otago, New Zealand

Outline

1. Introduction
2. Forward model
3. Bayesian inference
4. MCMC techniques
5. Simulation study
6. Conclusion and issues

Introduction

Basic scattering problem :
Concerns the effect of an inhomogeneous medium on an incident acoustic wave.

$$u \text{ (total field)} = u_i \text{ (incident field)} + u_s \text{ (scattered field)}$$

Inverse problem :
Determine the shape or physical properties of the obstacle from the measurement of u_s.

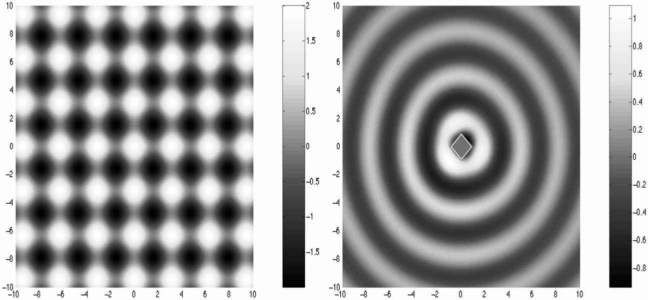


Figure: The real part of u_i (Left) and u_s (Right) for a diamond shaped obstacle

Direct Scattering Problem

Determine u_s from knowledge of u_i and the scattering obstacle.
Assuming the time harmonic waves

$$U(x,t) = u(x)e^{-iwt}$$

the u_s is estimated using the following equations,

$$\begin{aligned} \text{Helmholtz equation} &: \quad \nabla^2 u + k^2 u = 0 \ \text{ in } \Omega \in \mathbb{R}^2 \\ \text{Neumann B.C} &: \quad \frac{\partial u}{\partial n} = 0 \ \text{ on } \partial\Omega \\ \text{Radiation condition} &: \quad \lim_{r\to\infty} r\left(\frac{\partial u_s}{\partial r} - iku_s\right) = 0 \end{aligned}$$

where
k : wave number
n : the unit outward normal to $\partial\Omega$
$r = |x|$.

Direct Scattering Problem

Determine u_s from knowledge of u_i and the scattering obstacle.
Assuming the time harmonic waves

$$U(x,t) = u(x)e^{-iwt}$$

the u_s is estimated using the following equations,

$$\begin{aligned} \text{Helmholtz equation} \quad &: \quad \nabla^2 u + k^2 u = 0 \ \text{ in } \Omega \in \mathbb{R}^2 \\ \text{Neumann B.C} \quad &: \quad \frac{\partial u}{\partial n} = 0 \ \text{ on } \partial\Omega \\ \text{Radiation condition} \quad &: \quad \lim_{r\to\infty} r\left(\frac{\partial u_s}{\partial r} - iku_s\right) = 0 \end{aligned}$$

The Neumann boundary condition corresponds to a sound-hard obstacle and the Sommerfield radiation condition guarantees that u_s is outgoing.

Direct Scattering Problem

Determine u_s from knowledge of u_i and the scattering obstacle. Assuming the time harmonic waves

$$U(x,t) = u(x)e^{-iwt}$$

the u_s is estimated using the following equations,

$$\begin{aligned}
\text{Helmholtz equation} \quad &: \quad \nabla^2 u + k^2 u = 0 \ \text{ in } \Omega \in \mathbb{R}^2 \\
\text{Neumann B.C} \quad &: \quad \frac{\partial u}{\partial n} = 0 \ \text{ on } \partial\Omega \\
\text{Radiation condition} \quad &: \quad \lim_{r\to\infty} r\left(\frac{\partial u_s}{\partial r} - iku_s\right) = 0
\end{aligned}$$

- Also called the *forward map*, $\mathcal{M}_f$ since it maps the image of the obstacle to the scattering field (i.e., $\mathcal{M}_f$:Image$\rightarrow u_s$).

Direct Scattering Problem

- Solving a forward map for u_s becomes a exterior boundary value problem (BVP)
- Numerical and analytical solutions for the exterior Helmholtz equation.
- Analytical solution for a limited case.
- Numerical solution using the Green's function and Green's formulae

Exterior boundary value problem (BVP)

For a sufficiently smooth $\partial\Omega$, the solution for u_s is

$$\int_{\partial\Omega}\left[\frac{\partial u_i(x)}{\partial n}g(x|\xi)+\frac{\partial g(x|\xi)}{\partial n}u_s(x)\right]dl(x) \quad = \quad \begin{cases} u_s(\xi)\ ,\ \xi\in\Omega \\ \dfrac{u_s(\xi)}{2}\ ,\ \xi\in\partial\Omega \end{cases}$$

where

$$g(x|\xi)=\frac{i}{4}H_0(k|x-\xi|)\ , \qquad \frac{\partial g(x|\xi)}{\partial n}=-\frac{ik}{4}H_1(k|x-\xi|)\frac{\partial(|x-\xi|)}{\partial n}\ ,$$

H_0 : a Hankel function of the first kind of order zero
H_1 : a Hankel function of the first kind of order one

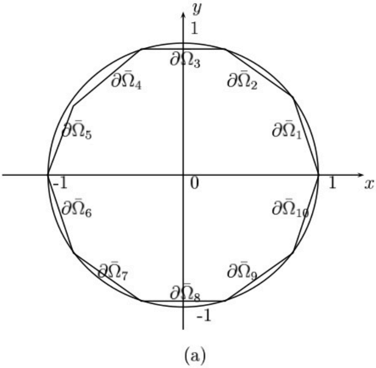
Exterior boundary value problem (BVP)

For a sufficiently smooth $\partial\Omega$, the solution for u_s is

$$\int_{\partial\Omega}\left[\frac{\partial u_i(x)}{\partial n}g(x|\xi)+\frac{\partial g(x|\xi)}{\partial n}u_s(x)\right]dl(x) = \begin{cases} u_s(\xi)\ ,\ \xi\in\Omega \\ \frac{u_s(\xi)}{2}\ ,\ \xi\in\partial\Omega \end{cases}$$

This is solved numerically using the boundary element method (BEM).
Boundary $\partial\Omega$ is discretized by N_b number of elements, $\{\partial\bar{\Omega}_i\}_{i=1}^{N_b}$
i.e., $\partial\Omega \approx \cup_{i=1}^{N_b}\partial\bar{\Omega}_i$

Figure: (a) the set of boundary elements $\{\partial\bar{\Omega}_i\}_{i=1}^{10}$ and the original boundary which is a unit circle and (b) the measurement points $s_1, \ldots, s_{10}$ around an unit circle obstacle.



Inverse Scattering Problem

- Unknown shape of the obstacle from the measurement of the scattering field u_s
- Nonlinear and ill-posed problem
- Small variations in u_s can lead to large errors in the reconstruction of the obstacle

Statistical Inference

Using a Bayesian approach, the inverse problem is tackled as a statistical inference for an unknown shape of the obstacle

- Γ - the continuous state space of feasible images w
- d - noisy measurements of the far field patterns

The posterior density for w is

$$p(w|d) = \frac{p(d|w)p(w)}{\int_\Gamma p(d|w)p(w)dw}$$

Statistical Inference

- *Synthetic Data* - u_s (synthetic noise free data) respect to the true image is obtained by solving the *forward map*. We assume both the real and imaginary parts of each measurement contain zero-mean Gaussian-distributed noise in reality.
- *Prior*,($p(w)$) - Since there is no real prior knowledge of w and hence no subjective prior. A uniform distribution over Γ is used.
- *Likelihood*, $p(d|w)$

$$p(d|w) = \prod_{i=1}^{N_f} \exp\left(\frac{-(\mathrm{Re}(d(s_i) - u_s(s_i)))^2 - (\mathrm{Im}(d(s_i) - u_s(s_i)))^2}{2\sigma^2}\right)$$

N_f : the number of measurements

Statistical Inference

The expected shape of obstacle is

$$\mathbb{E}[w] = \int_\Gamma wp(w|d)dw \ .$$

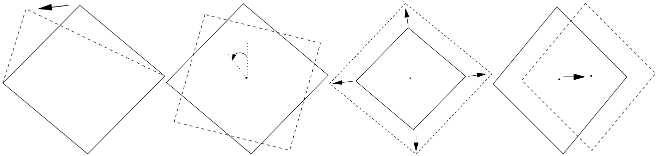
If the sample set $\{w_i\}_{i=1}^N$ generated from $p(w|d)$ over Γ the expectation of w is estimated by Monte Carlo integration

$$\bar{w} \approx \frac{1}{N} \sum_i w_i \ .$$

Markov chain Monte Carlo method is used to generated samples from $p(w|d)$, and the central limit theorem (CLT) holds for $\bar{w}$.

Markov chain Monte Carlo (MCMC) method

Figure: Four types of proposal movements; $\mathcal{V}$, $\mathcal{R}$, $\mathcal{S}$ and $\mathcal{T}$ (Left to Right)



MH-Algorithm

Given a state $W_n = w$, the Metropolis Hastings algorithm is as follows

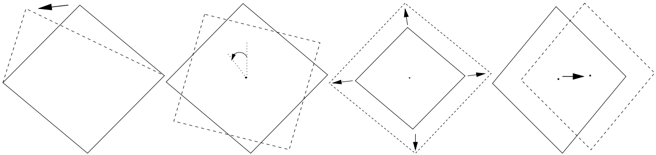
Step 1 Select a move m_i. $i \in \{1, \dots, N_m\}$ with probability ε_i. Generate w' by sampling $\Phi_i(w \to w')$.

Step 2 Compute the acceptance probability for the state, $\alpha^{m_i}(w \to w')$.

Step 3 Accept w' with $\alpha^{m_i}(w \to w')$.

Markov chain Monte Carlo (MCMC) method

Figure: Four types of proposal movements; $\mathcal{V}$, $\mathcal{R}$, $\mathcal{S}$ and $\mathcal{T}$ (Left to Right)



Move $\mathcal{R}$ Rotate w by a random angle $h^{\mathcal{R}}$ with respect to the center of mass, $h^{\mathcal{R}} \sim U(-\delta^{\mathcal{R}}, \delta^{\mathcal{R}})$

Move $\mathcal{T}$ Shift w by a random vector $h^{\mathcal{T}}$, $h^{\mathcal{T}} \sim U(-\delta^{\mathcal{T}}, \delta^{\mathcal{T}})^2$

Move $\mathcal{V}$ Move a position of one vertex by a random vector $h^{\mathcal{V}}$, $h^{\mathcal{V}} \sim U(-\delta^{\mathcal{V}}, \delta^{\mathcal{V}})^2$

Move $\mathcal{S}$ Change a size of w by a random rate $h^{\mathcal{S}}$ respect to the center of mass, $h^{\mathcal{S}} \sim U(\frac{1}{\delta^{\mathcal{S}}}, \delta^{\mathcal{S}})$

Computing the Likelihood efficiently

- Precomputation of Hankel function : Instead of evaluating a truncated infinite sum or similar for each Greens function value, look up the table of precomputed Hankel function values.
- Efficient boundary discretization : In solving the forward problem numerically, the number of elements N_b relates to the size of dimensions of the linear system.

Delayed acceptance MCMC (DAMCMC)

- Generates samples from the exact posterior using an intermediate approximation step.
- If a proposal is accepted by the approximation, it is corrected by calculating the true posterior density to ensure it reaches the target distribution. Otherwise a proposal is rejected.
- Computation time reduction by avoiding calculation of the exact density for proposals that are rejected.
- Quality of approximation.
- The speed up of these algorithms over the standard MH algorithm equals the inverse of the proportion of acceptance.

Delayed acceptance MCMC (DAMCMC)

Algorithm

Let $W_n = w$ and $p^*_w(w'|d)$ denotes the approximation to $p(w'|d)$ computed at w (i.e., $p^*_w(w|d) = p(w|d)$).

Step 1 Select a move m_i. $i \in \{1, \ldots, N_m\}$ with probability ε_i. Generate w' by sampling $\Phi_i(w \to w')$.

Step 2 Using the approximation of the present section, estimate $p^*_w(w'|d)$.

Step 3 Compute the Metropolis Hastings ratio $\alpha^{m_i}_{MH}(w \to w')$ using $p^*_w(w'|d)$.

Step 4 Accept or reject the candidate state w':

4.1 If w' is rejected, set $W_{n+1} = w$ and go back to **Step 1**.

4.2 If w' is accepted, compute $p_w(w'|d)$ using the exact calculation. Compute the second-stage acceptance probability $\alpha^{m_i}(w \to w')$ and accept w' with $\alpha^{m_i}(w \to w')$.

Forward map approximation

- Linear approximate
  Use of the the Fréchet derivative of the forward map.
- Coarse discretization approximate
  Use of a coarsened boundary discretization using fewer boundary elements than the exact calculation.

Forward map approximation

- Linear approximate (Hettlich (1995))
  For sufficiently small vector fields $h \in C^1(\partial\Omega)$ a perturbation of $\partial\Omega$ is a boundary of a domain denoted by $\partial\Omega_h = \{x' \in \mathbb{R}^2 : x' = x + h(x), x \in \partial\Omega\}$ in the class C^1. The normal component of a vector field h is denoted by $h_n = h \cdot n$ and the notation $(\nabla u)_t$ for the tangential component $(\nabla u)_t = n \times (\nabla u \times n)$ of a vector field u is used. Then the operator $\mathcal{M}$ is Fréchet differentiable at $\partial\Omega$, i.e.,

$$\frac{1}{\|h\|_{C^1(\partial\Omega)}} \|\mathcal{M}(\partial\Omega_h) - \mathcal{M}(\partial\Omega) - \mathcal{M}'(\partial\Omega)h\| \to 0 \ , \qquad h \to 0 \ .$$

  Note that the derivative is related to u'_∞, (i.e., $\mathcal{M}'(\partial\Omega)h = u'_\infty$). u'_∞ is the far-field pattern of the radiating solution of

$$\nabla^2 u'_\infty + k^2 u'_\infty = 0 \qquad \text{in } \Omega \subset \mathbb{R}^2$$

$$\frac{\partial u'_\infty}{\partial n} = k^2 h_n u_s + \mathrm{Div}(h_n (\nabla u_s)_t) \qquad \text{on } \partial\Omega \ .$$

  u_s denotes the solution of the scattering problem with respect to Ω. The linear approximation u'_s to the scattering field on a domain Ω_h for small h is given by

$$u'_s \simeq u_s + u'_\infty \ .$$

- Coarse discretization approximate

Forward map approximation

- Linear approximate
- Coarse discretization approximate
  - Fine discretization gives an accurate solution and demands a higher computation cost..
  - Coarse discretization using fewer boundary elements but also a reasonably accurate solution.
  - Separation analysis of coupled Markov chains (Nicholls, G. K., Fox, C. and Watt, A. (2012))
  - Key idea is to compare the mean separation time of two chains under the same random sequences and if the time is long enough, two chains are empirically identical.

- Linear approximate
- Coarse discretization approximate

  Algorithm

  Generate W_0 form the initial state distribution ρ_0 and set $Z_0 = W_0$.

  For $n < N_t$,

  Step 1 Generate a proposal w' from $\psi(w'|w)$.

  Step 2 Compute the acceptance probabilities; $\alpha_n^w(w'|w)$ using $p_w(w'|d)$ and $\alpha_n^z(w'|w)$ using $p_z(w'|d)$.

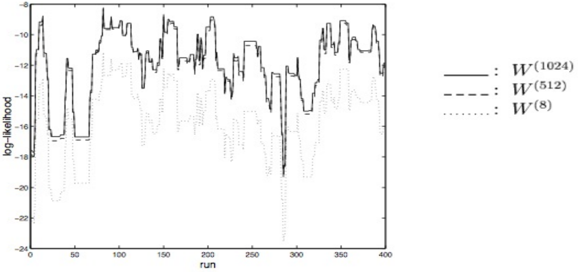
  Step 3 Accept w' with $\alpha_n^w(w'|w)$. If w' is accepted set $W_{n+1} = Z_{n+1} = w'$. Otherwise set $W_{n+1} = Z_{n+1} = w$.

  Step 4 After N_t iterations compute the mean absolute differences of α_w and α_z.

$$\alpha_{w,z} = \frac{1}{N_t} \sum_{n=1}^{N_t} |\alpha_n^w - \alpha_n^z|$$

- Linear approximate
- Coarse discretization approximate

Figure: log of likelihood of $W^{(1024)}$, $W^{(512)}$ and $W^{(8)}$

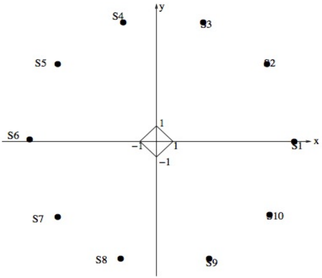


- Linear approximate
- Coarse discretization approximate
  - For a large N_t, $\alpha_{w,z}$ is the average separation probability per update and the inverse relates to the mean separation time of Z from W.
  - For theoretical result, see Nicholls, G. K., Fox, C. and Watt, A. M. (2012).
  - If a separation time is longer than the total run length, the approximate chain is identical to the exact chain.
  - If $\alpha_{w,z}$ is greater than an autocorrelation time of the exact, we treat that the chain mixes faster than it separates.

Simulation study

- A diamond shape of obstacle.
- Ten u_s measurements points, $(10\cos(\phi_i), 10\sin(\phi_i))$ where $\phi_i = 2\pi(i-1)/10$, $i = 1, \ldots, 10$.
- Two incident fields, $u_i^1 = e^{ikx}$ and $u_i^2 = e^{iky}$.
- Synthetic data : $d_i = u_s(s_i) + \epsilon_1 + i\epsilon_2$, $\epsilon_1, \epsilon_2 \sim N(0,\sigma^2)$.
- 512 boundary elements are used, $N_b = 512$.

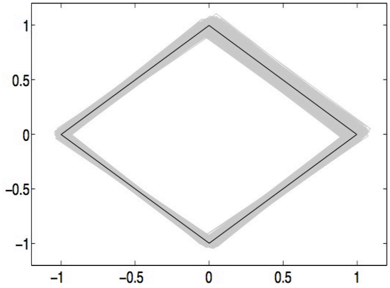
Figure: A diamond-shaped obstacle and the measurement points $s_1, \ldots, s_{10}$



Simulation study

- A diamond shape of obstacle.
- Ten u_s measurements points, $(10\cos(\phi_i), 10\sin(\phi_i))$ where $\phi_i = 2\pi(i-1)/10$, $i = 1, \ldots, 10$.
- Two incident fields, $u_i^1 = e^{ikx}$ and $u_i^2 = e^{iky}$.
- Synthetic data : $d_i = u_s(s_i) + \epsilon_1 + i\epsilon_2$, $\epsilon_1, \epsilon_2 \sim N(0, \sigma^2)$.
- 512 boundary elements, $N_b = 512$, for the exact chain.

Figure: Empirical posterior mean (black line) and 95% confidence intervals (grey lines).



Simulation study

$[\varepsilon_{\mathcal{R}}\ \varepsilon_{\mathcal{T}}\ \varepsilon_{\mathcal{V}}\ \varepsilon_{\mathcal{S}}]$	chain	L	τ	$\tau \times t$	N
[0 0 1 0]	W	-10.7925±0.0218	71.9718	24.3731	600000
[0 0 1 0]	W^C	-10.7914±0.0246	90.1811	17.6190	600000
[0 0 1 0]	W^L	-10.7932±0.0457	312.4799	94.7306	600000
[0.07 0.07 0.80 0.06]	W	-10.7894±0.0227	79.1486	21.0694	600000
[0.07 0.07 0.80 0.06]	W^C	-10.8025±0.0255	97.8979	15.7486	600000
[0.07 0.07 0.80 0.06]	W^L	-10.8143±0.0486	344.6875	101.2250	600000
[0.10 0.10 0.70 0.10]	W	-10.7964±0.0233	82.2308	21.7089	600000
[0.10 0.10 0.70 0.10]	W^C	-10.7690±0.0260	102.0964	15.2867	600000
[0.10 0.10 0.70 0.10]	W^L	-10.8098±0.0438	294.3497	77.254	600000
[0.20 0.20 0.55 0.05]	W	-10.7639±0.0268	108.5473	28.8844	600000
[0.20 0.20 0.55 0.05]	W^C	-10.7696±0.0283	121.4257	21.7538	600000
[0.20 0.20 0.55 0.05]	W^L	-10.7625±0.0450	313.8210	82.5361	600000

Summary

- The inverse obstacle scattering problem using Bayesian inference
- Two attempts to increase the forward map calculation efficiency ; precomputation of Hankel function values and efficient boundary discretization.
- The delayed acceptance MH algorithm can be effective in reducing a computational workload when an appropriate approximate is used.

Future works

- Advanced MCMC techniques such as adaptive delayed acceptance M-H algorithm by Cui, T., Fox, C., and O'Sullivan, M. J. (2011).
- Alternative representation for an obstacle shape.
- Improvement of forward map approximate.

References

Christen, A. and Fox, C. (2005). MCMC using an approximate. Journal of Computation and Graphical Statistics.

Cui, T., Fox, C. and O'Sullivan, M. J. (2011). Adaptive error modelling in MCMC sampling for large scale inverse problems. ISSN 1178-360. Report. Faculty of Engineering. No. 687, University of Auckland.

Hettlich, F. (1995). Frechet derivatives in inverse obstacle scattering. Inverse Problem, Vol 11. 371-382.

Lee, J. (2005). Sample Based Inference for Inverse Obstacle Scattering. MSc thesis. University of Auckland.

Nicholls, G. K., Fox, C. and Watt, A. (2012). Coupled MCMC with a randomized acceptance probability. (submitted). http://arxiv.org/abs/1205.6857.

Thank you for your attention!