Hierarchical Bayesian Level Set Inversion

Matt Dunlop

Mathematics Institute, University of Warwick, U.K.

Marco Iglesias (Nottingham), Andrew Stuart (Warwick)

SIAM UQ 2016,
Lausanne, Switzerland,
April 7th.

EPSRC

Outline

Outline

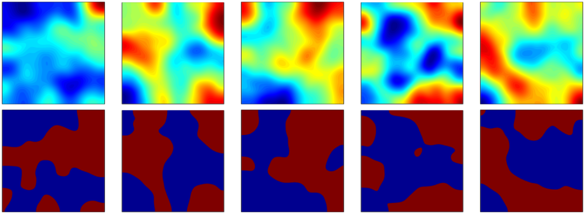
Level Set Representation

S.Osher and J.Sethian
Fronts propagating with curvature-dependent speed . . . .
J. Comp. Phys. **79**(1988), 12–49.

Piecewise constant function v defined through thresholding a continuous level set function u. Let

$$-\infty = c_0 < c_1 < \cdots < c_{K-1} < c_K = \infty.$$

$$v(x) = \sum_{k=1}^{K} v_k \, \mathbb{I}_{\{c_{k-1} < u \le c_k\}}(x); \quad v = F(u).$$



$F : X \to Z$ is the level-set map, X cts fns. Z piecewise cts fns.

Example 1: Image Reconstruction

H.W. Engl, M. Hanke and A. Neubauer
Regularization of Inverse Problems
Kluwer(1994)

Forward Problem

Define $K : Z \to \mathbb{R}^J$ by $(Kv)_j = v(x_j), x_j \in D \subset \mathbb{R}^d$. Given $v \in Z$

$$y = Kv.$$

Let $\eta \in \mathbb{R}^J$ be a realization of an observational noise.

Inverse Problem

Given prior information $v = F(u)$, $u \in X$ and $y \in \mathbb{R}^J$, find v :

$$y = Kv + \eta.$$

Example 2: Groundwater Flow

M. Dashti and A.M. Stuart
The Bayesian approach to inverse problems.
Handbook of Uncertainty Quantification
Editors: R. Ghanem, D.Higdon and H. Owhadi, Springer, 2017.
arXiv:1302.6989

Forward Problem: Darcy Flow

Let $X^+ := \{v \in Z : \mathrm{essinf}_{x\in D} v > 0\}$. Given $\kappa \in X^+$, find $y := \mathcal{G}(\kappa) \in \mathbb{R}^J$ where $y_j = \ell_j(p)$, $V := H_0^1(D), \ell_j \in V^*$, $j = 1, \dots, J$ and

$$\begin{cases} -\nabla \cdot (\kappa \nabla p) = f & \text{in } D \\ p = 0 & \text{on } \partial D. \end{cases}$$

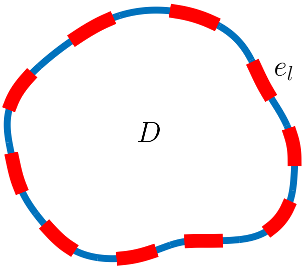
Let $\eta \in \mathbb{R}^J$ be a realization of an observational noise.

Inverse Problem

Given prior information $\kappa = F(u)$, $u \in X$ and $y \in \mathbb{R}^J$, find κ :

$$y = \mathcal{G}(\kappa) + \eta.$$

Example 3: Electrical Impedance Tomography



- Apply currents I_ℓ on e_ℓ, $\ell = 1, \ldots, L$.
- Induces voltages Θ_ℓ on e_ℓ, $\ell = 1, \ldots, L$.
- Input is (σ, I), output is (θ, Θ).
- We have an Ohm's law $\Theta = R(\sigma)I$.

$$\begin{cases} -\nabla \cdot (\sigma(x)\nabla\theta(x)) = 0 & x \in D \\ \displaystyle\int_{e_\ell} \sigma \frac{\partial\theta}{\partial\nu}\,\mathrm{d}S = I_\ell & \ell = 1, \ldots, L \\ \sigma(x)\dfrac{\partial\theta}{\partial\nu}(x) = 0 & x \in \partial D \setminus \bigcup_{\ell=1}^{L} e_\ell \\ \theta(x) + z_\ell\sigma(x)\dfrac{\partial\theta}{\partial\nu}(x) = \Theta_\ell & x \in e_\ell, \ell = 1, \ldots, L \end{cases} \tag{PDE}$$

Example 3: Electrical Impedance Tomography

M. M. Dunlop and A. M. Stuart
The Bayesian formulation of EIT.
arXiv:1509.03136
Inverse Problems and Imaging, Submitted, 2015.

Forward Problem

- Let $X \subseteq L^\infty(D)$, and denote $X^+ := \{u \in X : \operatorname{essinf}_{x \in D} u > 0\}$.
- Given $u \in X, F : X \to X^+$ and $\sigma = F(u)$, find $(\theta, \Theta) \in H^1(D) \times \mathbb{R}^L$ solving (PDE).
- This gives $\Theta = R(F(u))I$.

Let $\eta \in \mathbb{R}^L$ be a realisation of an observational noise.

Inverse Problem

Given prior information $\sigma = F(u)$, $u \in X$, and given currents I and $y \in \mathbb{R}^L$, find σ :

$$y = R(\sigma)I + \eta.$$

Outline

Tikhonov Regularization

$\mathcal{G} : Z \to \mathbb{R}^J$ and $\eta \in \mathbb{R}^J$ a realization of an observational noise.

Inverse Problem

Find $u \in X$, given $y \in \mathbb{R}^J$ satisfying $y = \mathcal{G} \circ F(u) + \eta$.

Model-Data Misfit

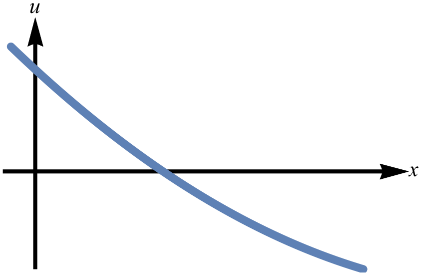
The model-data misfit is $\Phi(u; y) = \frac{1}{2}\left|\Gamma^{-\frac{1}{2}}\big(y - \mathcal{G} \circ F(u)\big)\right|^2$.

Tikhonov Regularization

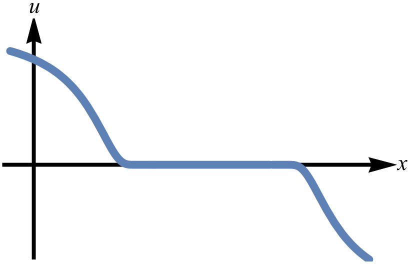
Minimize, for some E compactly embedded into X,

$$I(u; y) := \Phi(u; y) + \frac{1}{2}\|u\|_E^2.$$

Issue 1: Discontinuity of Level Set Map



$F(\cdot)$ is continuous at u

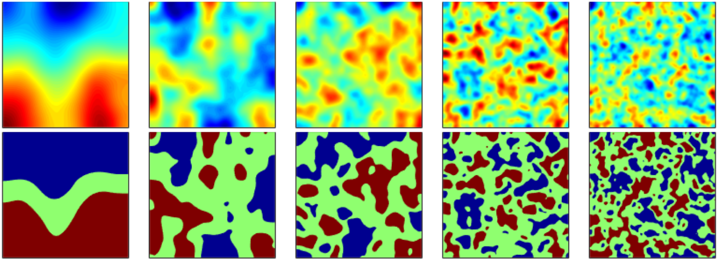


$F(\cdot)$ is discontinuous at u

$$v = F(u) := v^{+}\,\mathbb{I}_{\{u \geq 0\}}(x) + v^{-}\,\mathbb{I}_{\{u < 0\}}(x).$$

Causes problems in classical level set inversion.

Issue 2: Length-Scale Matters



- Figure demonstrates role of length-scale in level-set function.
- Classical Tikhonov-Phillips regularization does not allow for direct control of the length-scale.

New ideas needed.

Issue 3: Amplitude Matters

Y. Lu
Probabilistic Analysis of Interface Problems
PhD Thesis (2016), Warwick University

- Consider case of $K = 2$, $c_1 = 0$.
- For contradiction assume $u^\star$ is a minimizer of $I(u; y)$.
- Now define the sequence $u_\epsilon = \epsilon u^\star$. Then if $0 < \epsilon < 1$,

$$\Phi(u_\epsilon; y) = \Phi(u^\star; y), \|u_\epsilon\|_E = \epsilon\|u^\star\|_E \Rightarrow I(u_\epsilon; y) < I(u^\star; y).$$

- Hence $I(u_\epsilon; y)$ is an infimizing sequence. But $u_\epsilon \to 0$ and $I(0; y) > I(u_\epsilon; y)$ for $\epsilon \ll 1$.
- Thus infimum cannot be attained.
- This issue caused by thresholding at 0. But idea generalizes.

Choice of threshold matters.

Outline

Bayes' Formula

Prior

Probabilistic information about u before data is collected:

$$\mu_0(du)$$

Likelihood

Since $y = \mathcal{G}(u) + \eta$, if $\eta \sim N(0, \Gamma)$, then $y|u \sim N(\mathcal{G}(u), \Gamma)$. The model-data misfit Φ is the negative log-likelihood:

$$\mathbb{P}(y|u) \propto \exp\big(-\Phi(u;y)\big), \quad \Phi(u;y) = \frac{1}{2}\Big|\Gamma^{-\frac{1}{2}}\big(y - \mathcal{G}(u)\big)\Big|^2.$$

Posterior

Probabilistic information about u after data is collected:

$$\mu^y(du) \propto \exp\big(-\Phi(u;y)\big)\mu_0(du).$$

Rigorous Statement

M. Iglesias, Y. Lu and A. M . Stuart
A level-set approach to Bayesian geometric inverse problems.
arXiv:1504.00313

M. Iglesias
A regularizing iterative ensemble Kalman method for PDE constrained inverse problems.
arXiv:1505.03876

$$L^2_\nu(X;S) = \{f : X \to S : \mathbb{E}^\nu \|f(u)\|_S^2 < \infty\}.$$

Theorem (Iglesias, Lu and S)

Let $\mu_0(du) = \mathbb{P}(du)$ and $\mu^y(du) = \mathbb{P}(du|y)$. Assume that $u \in X$ μ_0-a.s. Then for Examples 1–3 (and more) $\mu^y \ll \mu_0$ and $y \mapsto \mu^y(du)$ is Lipschitz in the Hellinger metric. Furthermore, if S is a separable Banach space and $f \in L^2_{\mu_0}(X;S)$, then

$$\left\|\mathbb{E}^{\mu^{y_1}} f(u) - \mathbb{E}^{\mu^{y_2}} f(u)\right\|_S \le C|y_1 - y_2|.$$

Key idea in proof is that $F : X \to Z$ is continuous μ_0- a.s.

Whittle-Matern Priors

$$c(x,y) = \sigma^2 \frac{1}{2^{\nu-1}\Gamma(\nu)}(\tau|x-y|)^{\nu} K_{\nu}(\tau|x-y|).$$

- ν controls smoothness – draws from Gaussian fields with this covariance have ν fractional Sobolev and Hölder derivatives.
- τ is an inverse length-scale.
- σ is an amplitude scale.

Ignoring boundary conditions, corresponding covariance operator given by

$$C_{\mathrm{WM}(\sigma,\tau)} \propto \sigma^2 \tau^{2\nu} (\tau^2 I - \triangle)^{-\nu - \frac{d}{2}}.$$

Example 2: Groundwater Flow

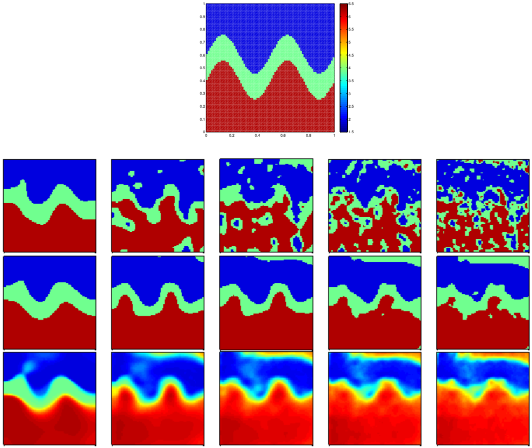


Figure: (Row 1, $\tau = 15$) Logarithm of the true hydraulic conductivity. (Row 2 $\tau = 10, 30, 50, 70, 90$) samples of $F(u)$. (Row 3) $F(\mathbb{E}(u))$. (Row 4) $\mathbb{E}(F(u))$.

Outline

Hierarchical Whittle-Matern Priors (The Problem)

Prior:

$$\mathbb{P}(u, \tau) = \mathbb{P}(u|\tau)\mathbb{P}(\tau).$$

$$\mathbb{P}(u|\tau) = N(0, C_{\mathrm{WM}(\sigma,\tau)}).$$

Recall $C_{\mathrm{WM}(\sigma,\tau)} \propto \sigma^2 \tau^{2\nu} (\tau^2 I - \triangle)^{-\nu - \frac{d}{2}}$.

Theorem

For fixed σ the family of measures $N(0, C_{\mathrm{WM}(\sigma,\cdot)})$ are mutually singular.

Algorithms which attempt to move between singular measures perform badly under mesh refinement (since they fail completely in the fully resolved limit).

Hierarchical Whittle-Matern Priors (The Solution)

M. Dunlop, M. Iglesias and A. M . Stuart
Hierarchical Bayesian Level Set Inversion
arXiv:1601.03605

Hence choose $\sigma = \tau^{-\nu}$ and define: $C_\tau \propto (\tau^2 I - \triangle)^{-\nu - \frac{d}{2}}$. Prior:

$$\mathbb{P}(u, \tau) = \mathbb{P}(u|\tau)\mathbb{P}(\tau).$$

$$\mathbb{P}(u|\tau) = N(0, C_\tau).$$

Theorem

The family of measures $N(0, C_\tau)$ are mutually equivalent.

Suggests need to scale thresholds in F by τ. Let

$$-\infty = c_0 < c_1 < \cdots < c_{K-1} < c_K = \infty.$$

$$v(x) = \sum_{k=1}^{K} v_k \, \mathbb{I}_{\{c_{k-1} < u\tau^\nu \le c_k\}}(x); \quad v = F(u, \tau).$$

Hierarchical Posterior

Prior

$$\mu_0(\mathrm{d}u, \mathrm{d}\tau) = \mu_0(\mathrm{d}u \,|\, \tau)\pi_0(\mathrm{d}\tau)$$

where $\mu_0(\,\cdot\,|\tau) = N(0, C_\tau)$, π_0 is hyperprior on τ.

Likelihood

Likelihood is given by $\mathbb{P}(y|u,\tau) \propto \exp\Big(-\Phi(u,\tau;y)\Big)$. where the model-data misfit is now $\Phi(u,\tau;y) = \frac{1}{2}\Big|\Gamma^{-\frac{1}{2}}\big(y - \mathcal{G}\circ F(u,\tau)\big)\Big|^2$.

Posterior

Probabilistic information about u after data is collected:

$$\mu^y(\mathrm{d}u, \mathrm{d}\tau) \propto \exp\big(-\Phi(u,\tau;y)\big)\mu_0(\mathrm{d}u, \mathrm{d}\tau).$$

MCMC

We implement a Metropolis-within-Gibbs algorithm to generate a posterior-invariant Markov chain (u_k, τ_k):

- Propose v_{k+1} from a $\mathbb{P}(u|\tau_k)-$reversible kernel.
- $u_{k+1} = v_{k+1}$ w.p. $1 \wedge \exp(\Phi(u_k, \tau_k) - \Phi(v_{k+1}, \tau_k))$,
  $u_{k+1} = u_k$ otherwise.
- Propose t_{k+1} from a $\mathbb{P}(\tau)-$reversible kernel.
- $\tau_{k+1} = t_{k+1}$ w.p. $1 \wedge \exp(\Phi(u_{k+1}, \tau_k) - \Phi(u_{k+1}, t_{k+1}))w(\tau_k, t_{k+1})$,
  $\tau_{k+1} = \tau_k$ otherwise.

Here $w(\tau, t)$ is density of $N(0, C_t)$ with respect to $N(0, C_\tau)$.

Example 2: Groundwater Flow

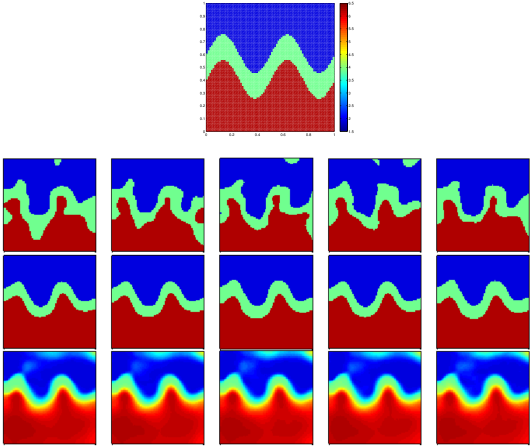


Figure: (Row 1, $\tau = 15$) Logarithm of the true hydraulic conductivity (middle, top). (Row 2 $\tau = 10, 30, 50, 70, 90$) samples of $F(u, \tau)$. (Row 3) $F(\mathbb{E}(u), \mathbb{E}(\tau))$. (Row 4) $\mathbb{E}(F(u, \tau))$.

Outline

Summary

- Last 5 years development of a theoretical and computational framework for infinite dimensional Bayesian inversion with wide applicability:
  1. Framework allows for noisy data and uncertain prior information.
  2. Probabilistic well-posedness.
  3. Theory clearly delineates (and links) analysis and probability.
  4. Theory leads to new algorithms (defined on Banach space).
  5. Grid-independent convergence rates for MCMC.
- The methodology has been extended to solve interface problems. This is achieved via the level set representation.
  1. Level set method becomes well-posed in this setting.
  2. Discontinuity in level set map is a probability zero event.
  3. Hierarchical choice of length-scale improves performance.
  4. Amplitude scale is linked to length-scale via measure equivalence.
  5. Algorithms which reflect this link are mesh-invariant.
- Potential for many future developments for both applications and theory:
  1. Applications: subsurface imaging, medical imaging.
  2. Theory: inhomogenous length-scales, new hierarchies.