Bayesian Inverse Problems in PDEs

Masoumeh Dashti and Andrew Stuart[1]

[1]Mathematics Institute and
Centre for Scientific Computing
University of Warwick

ICMS, Uncertainty Quantification
May 26$^{\mathrm{th}}$ 2010

Funded by EPSRC, ERC, ONR

Outline

Outline

The Idea

- WELL-POSED $\oplus$ SMALL PERTURBATION TO EQUATION $\Rightarrow$ SMALL PERTURBATION TO SOLUTION
- STABILITY $\oplus$ CONSISTENCY $\Rightarrow$ CONVERGENCE
- WELL-POSED $\oplus$ APPROXIMATION OF FORWARD PROBLEM $\Rightarrow$ APPROXIMATION OF INVERSE PROBLEM

Outline

Groundwater Flow

- Let $\mathcal{H} = L^2(D, \mathbb{R}^3)$.
- Consider Darcy's Law

$$\nabla \cdot \Big(\exp(u)\nabla p\Big) = 0, \quad x \in D$$
$$p = \phi, \quad x \in \partial D.$$

- Find u.
- Given noisy *observations* of the pressure:

$$y_j = p(x_j) + \eta_j$$

Outline

Least Squares

- The equation $y = \mathcal{G}(u)$ for $u \in X$ is ill-posed: (no solution, many solutions, sensitive dependence).
- Replace by least squares problem

$$\min_{u \in X} \Phi(u; y), \quad \Phi(u; y) = \frac{1}{2}\|y - \mathcal{G}(u)\|_Y^2.$$

- Can have $\Phi(u_n; y) \to 0$ but u_n does not converge in X.
- Tikhonov Regularization:

$$\min_{u \in E} I(u; y), \quad I(u; y) = \Phi(u; y) + \frac{1}{2}\|u\|_E^2.$$

Outline

Bayesian Approach to Inverse Problems

- Jointly varying random variable $(u, y) \in X \times Y$.
- u prior $\mu_0 = \mathbb{P}(du)$ on u :

$$\mu_0 = \mathcal{N}(0, \mathcal{C}_0), \quad \mu_0(X) = 1.$$

- y|u data $y \in Y$

$$y = \mathcal{G}(u) + \eta, \quad \eta \sim \mathcal{N}(0, \Gamma).$$

- u|y posterior $\mu^y = \mathbb{P}(du|y)$ on u :

$$\frac{d\mu^y}{d\mu_0}(u) \propto \mathbb{P}(y|u) \propto \exp\Big(-\Phi(u; y)\Big).$$

Potential

Conditions on the Potential

- there is, $\forall \epsilon > 0, r > 0$ an $M = M(\epsilon, r) > 0$ such that

$$\Phi(u; y) \geq -\epsilon \|u\|_X^2 + M \quad \forall u \in X, \|y\|_Y < r;$$

- $\forall r > 0$ there is $K(r) > 0$ such that,

$$|\Phi(u_1; y) - \Phi(u_2; y)| \leq K(r)\|u_1 - u_2\|_X, \quad \forall \|u_i\|_X, \|y\|_Y < r;$$

- $\forall \epsilon > 0, r > 0$, there is $K = K(\epsilon, r) > 0$ such that,

$$|\Phi(u; y_1) - \Phi(u; y_2)| \leq K e^{\epsilon \|u\|_X^2} \|y_1 - y_2\|_Y, \; \forall u \in X, \|y_i\|_Y < r.$$

Well-Posed Inverse Problem

Theorem

Assume that POTENTIAL CONDITIONS *hold and that* $\mu_0(X) = 1$. *There is* $C = C(r) > 0$ *such that, for all* y_1, y_2 *with* $\max\{\|y_1\|_Y, \|y_2\|_Y\} \le r$,

$$d_{Hell}(\mu^{y_1}, \mu^{y_2}) \le C\|y_1 - y_2\|_Y.$$

The metric d_{Hell} is a useful one because:

$$\|\mathbb{E}^{\mu} f - \mathbb{E}^{\nu} f\| \le 2\Big(\mathbb{E}^{\mu}\|f\|^2 + \mathbb{E}^{\nu}\|f\|^2\Big)^{\frac{1}{2}} d_{Hell}(\mu, \nu).$$

Probability Maximizers and Optimal Control

Probability maximizers are minimizers of the least squares functional

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

And this minimization is well-defined:

Theorem

Assume that POTENTIAL CONDITIONS *hold and that* $\mu_0(X) = 1$. *Then there exists* $\overline{u} \in E$ *such that*

$$I(\overline{u}) = \bar{I} := \inf\{I(u) : u \in E\}.$$

Furthermore, if $\{u_n\}$ *is a minimizing sequence satisfying* $I(u_n) \to I(\overline{u})$ *then there is a subsequence* $\{u_{n'}\}$ *that converges strongly to* $\overline{u}$ *in* E.

Forward Approximation Gives Inverse Approximation

Consider two measures μ and μ^N :

$$\frac{d\mu}{d\mu_0}(u) \propto \exp\Big(-\Phi(u)\Big), \quad \frac{d\mu^N}{d\mu_0}(u) \propto \exp\Big(-\Phi^N(u)\Big).$$

Theorem

Assume that: Φ *and* Φ^N *satisfy* POTENTIAL CONDITIONS *uniformly in* N *and that* $\mu_0(X) = 1$. *Assume also that, for any* $\epsilon > 0$ *there is* $K = K(\epsilon) > 0$ *such that*

$$|\Phi(u) - \Phi^N(u)| \leq K \exp\big(\epsilon \|u\|_X^2\big)\psi(N) \tag{1}$$

where $\psi(N) \to 0$ *as* $N \to \infty$.
Then there is a constant C*, independent of* N*, and such that*

$$d_{Hell}(\mu, \mu^N) \leq C\psi(N). \tag{2}$$

Outline

Problem Statement

- Consider Darcy's Law

$$\nabla \cdot \Big(\exp(u)\nabla p\Big) = 0, \quad x \in D$$
$$p = \phi, \quad x \in \partial D.$$

- For any $u \in L^\infty(D)$ define:
- $\lambda(u) = \operatorname{ess\,inf}_{x\in D} \mathrm{e}^{u(x)} > 0$;
- $\Lambda(u) = \operatorname{ess\,sup}_{x\in D} \mathrm{e}^{u(x)} < \infty$.
- We do not assume that the upper and lower bounds on λ/Λ hold uniformly across the probability space.

Elliptic problem: Observations

Noisy observations of p at a set of points $x_1, \ldots, x_K$ in D:

$$y_k = p(x_k) + \eta_k, \quad k = 1, \ldots, K$$

We assume that: the noise is Gaussian
$\{\eta_k\}$ an i.i.d sequence with $\eta_1 \sim \mathcal{N}(0, \gamma^2 I)$

Concatenating the data, we have

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

$$\begin{aligned} y &= (y_1, \cdots, y_K)^T; \\ \eta &= (\eta_1, \ldots, \eta_K)^T \sim \mathcal{N}(0, \gamma^2 I) \\ \mathcal{G}(u) &= (p(x_1), \cdots, p(x_K))^T, \quad \text{the observation operator.} \end{aligned}$$

Elliptic problem: Estimates

One can show that:

- if $u \in L^\infty(D)$ there exists $C = C(D, \|\phi\|_{L^\infty(\partial D)})$ such that

$$|\mathcal{G}(u)| \leq C\,\mathrm{e}^{\|u\|_{L^\infty(D)}}$$

- if $u_1, u_2 \in C^\alpha(D)$ for some $\alpha > 0$ then there exists $C = C(D, \alpha, \epsilon)$ such that

$$|\mathcal{G}(u_1) - \mathcal{G}(u_2)| \\ \leq C \exp\left(\epsilon \max\{\|u_1\|^2_{C^\alpha(D)}, \|u_2\|^2_{C^\alpha(D)}\}\right) \|u_1 - u_2\|_{L^\infty(D)}$$

- Hence $\Phi = \frac{1}{2}|y - \mathcal{G}(u)|^2$ satisfies potential conditions.

Elliptic inverse problem: Well-posedness of the posterior

Therefore

if $\mu_0(C^\alpha(D)) = 1$, μ^y is absolutely continuous with Radon-Nikodym derivative

$$\frac{\mathrm{d}\mu^y}{\mathrm{d}\mu_0}(u) \propto \exp\Big(-\frac{1}{2\gamma^2}|y - \mathcal{G}(u)|^2\Big)$$

and

$$d_{Hell}(\mu^y, \mu^{y'}) \leq C\,|y - y'|$$

for any $y, y' \in \mathbb{R}^K$.

Need a μ_0 that satisfies $\mu_0(C^\alpha(D)) = 1$

Elliptic inverse problem: Prior

Let $\mu_0 = \mathcal{N}(0, \mathcal{C})$ with $\mathcal{C} : L^2(D) \to L^2(D)$,

$\{\lambda_k\}_{k\in\mathbb{N}}$ and $\{\phi_k\}_{k\in\mathbb{N}}$ eigenvalues and eigenvectors of $\mathcal{C}$.

Then by the Karhunen-Loève expansion

$$u = \sum_k \sqrt{\lambda_k}\, \phi_k\, \xi_k \quad \text{with } \xi_k \text{ i.i.d and } \xi_1 \sim \mathcal{N}(0,1).$$

Let $\mathcal{A} = -\Delta$ (Δ: the Laplacian operator) acting on

$$D(\mathcal{A}) = \left\{ u \in H^2(D) : \nabla u \cdot \mathbf{n} = 0 \text{ on } \partial D, \text{ and } \int_D u = 0 \right\}.$$

Let $\mu_0 = \mathcal{N}(u_b, k\mathcal{A}^{-\beta})$, $k > 0$, $u_b \in H^\beta$.
If $\beta > d - 1/2$, then

$$\mu_0\left(C^\alpha(\overline{D})\right) = 1 \quad \text{for any } \alpha < \frac{1}{3}(\beta - (d - 1/2)).$$

Elliptic inverse problem: Prior

Let $\mu_0 = \mathcal{N}(0, \mathcal{C})$ with $\mathcal{C} : L^2(D) \to L^2(D)$,

$\{\lambda_k\}_{k\in\mathbb{N}}$ and $\{\phi_k\}_{k\in\mathbb{N}}$ eigenvalues and eigenvectors of $\mathcal{C}$.

Then by the Karhunen-Loève expansion

$$u = \textstyle\sum_k \sqrt{\lambda_k}\, \phi_k\, \xi_k \quad \text{with } \xi_k \text{ i.i.d and } \xi_1 \sim \mathcal{N}(0,1).$$

Let $\mathcal{A} = -\Delta$ (Δ: the Laplacian operator) acting on

$$D(\mathcal{A}) = \left\{ u \in H^2(D) : \nabla u \cdot \mathbf{n} = 0 \text{ on } \partial D, \;\; \text{and} \;\; \int_D u = 0 \right\}.$$

Let $\mu_0 = \mathcal{N}(u_b, k\mathcal{A}^{-\beta})$, $k > 0$, $u_b \in H^\beta$.
If $\beta > d - 1/2$, then

$$\mu_0\left(C^\alpha(\overline{D})\right) = 1 \quad \text{for any } \; \alpha < \frac{1}{3}(\beta - (d - 1/2)).$$

Elliptic inverse problem: Prior

Let μ_0 be distributed as $\mathcal{N}(0, \mathcal{C})$ with

$$\mathcal{C}u(x) = \int_D c(x, y)\, u(y)\, \mathrm{d}y, \qquad c(x, y) = \mathrm{e}^{-|x-y|}, \quad \text{or} \quad \mathrm{e}^{-\sum_{j=1}^{d} |x_j - y_j|}$$

Then, for any $\alpha < \frac{1}{2}$, $\mu_0\left(C^\alpha(D)\right) = 1$.

Elliptic inverse problem: Approximations

- $\{\psi_{j,k} = \mathrm{e}^{i(jx+ky)}\}_{j,k\in\mathbb{Z}}$ basis for $L^2(\mathbb{T}^2)$.
- $W^N = \mathrm{span}\{\psi_{j,k}, |j| \le N, |k| \le N\}$
- P^N the orthogonal projection of $L^2(D)$ onto W^N.

Let $u^N = P^N u \;\; \mathcal{G}^N(\cdot) = \mathcal{G}(P^N\cdot)$ and $\Phi^N(\cdot) = \Phi(P^N\cdot)$.

Define $\mu^{y,N}$ by

$$\frac{\mathrm{d}\mu^{y,N}}{\mathrm{d}\mu_0}(u) = \frac{1}{Z^N}\exp\Big(-\frac{1}{2\gamma^2}|y-\mathcal{G}^N(u)|^2\Big)$$

with $\quad Z^N = \int_X \exp\big(-\frac{1}{2\gamma^2}|y-\mathcal{G}^N(u)|^2\big)\mathrm{d}\mu_0(u).$

Elliptic inverse problem: Approximations

- $\{\psi_{j,k} = \mathrm{e}^{i(jx+ky)}\}_{j,k\in\mathbb{Z}}$ basis for $L^2(\mathbb{T}^2)$.
- $W^N = \mathrm{span}\{\psi_{j,k}, |j| \le N, |k| \le N\}$
- P^N the orthogonal projection of $L^2(D)$ onto W^N.

Let $u^N = P^N u$ $\mathcal{G}^N(\cdot) = \mathcal{G}(P^N\cdot)$ and $\Phi^N(\cdot) = \Phi(P^N\cdot)$.

Define $\mu^{y,N}$ by

$$\frac{\mathrm{d}\mu^{y,N}}{\mathrm{d}\mu_0}(u) = \frac{1}{Z^N}\exp\Big(-\frac{1}{2\gamma^2}|y - \mathcal{G}^N(u)|^2\Big)$$

with $\quad Z^N = \int_X \exp\big(-\frac{1}{2\gamma^2}|y - \mathcal{G}^N(u)|^2\big)\mathrm{d}\mu_0(u).$

Elliptic inverse problem: Approximations

- Choose μ_0 such that $\mu_0(C^\alpha(D)) = 1$ for some $\alpha > 0$.
- We have

$$|\Phi^N(u) - \Phi(u)| \leq C \exp\left(\epsilon \|u\|^2_{C^\alpha(D)}\right) N^{-\alpha} (\log N)^2.$$

Therefore

$$d_{Hell}(\mu^y, \mu^{y,N}) \leq C\, N^{-\alpha} (\log N)^2.$$

Outline

What We Have Shown

We have shown that:

- **Applications:** Inverse problems in differential equations can be formulated in the framework of Bayesian statistics on function space.
- **Common Structure:** Many problems share a common mathematical structure leading to *well-posed* inverse problems for measures and a link to *optimal control*.
- **Approximation:** This well-posedness leads to a transfer of approximation properties from the forward problem to the inverse problem, in the Hellinger metric.

References

- A.M. Stuart. "Inverse Problems: A Bayesian Perspective." Acta Numerica 2010.
- S.L.Cotter, M. Dashti, J.C.Robinson, A.M.Stuart. "Bayesian inverse problems for functions and applications to fluid mechanics." Inverse Problems, 25(2009), doi:10.1088/0266-5611/25/11/115008.
- S.L. Cotter, M. Dashti and A.M.Stuart. "Approximation of Bayesian Inverse Problems". Submitted, SIAM J. Num. Anal. 48(2010),322–345.

References (Continued)

- A. Beskos and A.M. Stuart. ”MCMC Methods for Sampling Function Space”. Invited Lectures, ICIAM 2007, Published by the European Mathematical Society, 2009.
- A. Beskos and A.M. Stuart. “Computational complexity of Metropolis-Hastings meth- ods in high dimensions”. Plenary Lectures MCQMC2008, Springer, 2009.
- A. Beskos, G.O. Roberts and A.M. Stuart. ”Optimal scalings for local Metropolis-Hastings chains on non-product targets in high dimensions.” Ann. Appl. Prob. **19**(2009), 863–898.

References (Continued)

- M. Hairer, A.M.Stuart and J. Voss. ”Sampling the posterior: an approach to non-Gaussian data assimilation.” PhysicaD, **230**(2007), 50–64.
- M. Hairer, A.M.Stuart, P. Wiberg and J. Voss. ”Analysis of SPDEs Arising in Path Sampling. Part 1: The Gaussian Case.” Comm. Math. Sci. 3(2005), 587–603
- M. Hairer, A.M.Stuart and J. Voss. ”Analysis of SPDEs Arising in Path Sampling. Part 2: The Nonlinear Case.” Ann. Appl. Prob. 17(2007), 1657–1706.
- For all papers see:

  $http://www.maths.warwick.ac.uk/\sim masdr/sample.html$