A Bayesian Approach to the Laplacian-like Inverse Problem

Sergios Agapiou Stig Larsson Andrew Stuart

The University of Warwick, Coventry, UK

ICIAM 2011

Summary

Inverse Problems are concerned with determining causes for a desired or an observed effect. In our context, we want to invert a linear operator from noisy observations of its image.

Aim: To study the consistency of the Bayesian Approach to Linear Inverse Problems in the small noise limit.

We start with the Laplacian-like Inverse Problem with noise and prior covariances diagonalizable in the eigenbase of the Laplacian and generalize to more general situations. We are always assuming a separable Hilbert-Space setting.

Numerical results are presented which support this theory.

Problem set-up - General Case

Consider the Inverse Problem to find u from y, for

$$y = \mathcal{A}^{-\ell}u + \frac{1}{\sqrt{n}}\xi, \ \ell > 0, \qquad (IP)$$

where $\xi \sim \mathcal{N}(0, C_1)$ and $\mathcal{A}$ is selfadjoint.

- Likelihood: for fixed u the law of $y|u$, called the *likelihood*, is $\mathcal{N}(\mathcal{A}^{-\ell}u, C_1)$.

- Prior: we now need to choose a *prior* distribution for the unknown u, based on any prior knowledge we have about u. Let $u \sim \mathcal{N}(0, \tau^2 C_0)$.

- Posterior: in the Bayesian Approach, the solution of (IP) is the distribution of $u|y$, called the *posterior* distribution. Here the posterior is $\mathcal{N}(m, C)$ where

$$m = C_0\mathcal{A}^{-\ell}(n^{-1}\tau^{-2}C_1 + \mathcal{A}^{-\ell}C_0\mathcal{A}^{-\ell})^{-1}y := Ly$$

and

$$C = \tau^2 C_0 - \tau^2 L\mathcal{A}^{-\ell}C_0.$$

Assume we have observations of the form

$$y_n = \mathcal{A}^{-\ell}u^\dagger + \frac{1}{\sqrt{n}}\xi \qquad (DM)$$

where $u^\dagger$ is the true solution. Then the posterior mean is

$$m_n = Ly_n. \qquad (PM)$$

We examine the consistency of the posterior mean in the sense

$$\mathbb{E}\left|\left|m_n - u^\dagger\right|\right|^2 \to 0$$

as $n \to \infty$ and for appropriate choice of $\tau = \tau(n)$.

Problem set-up - Diagonal Case

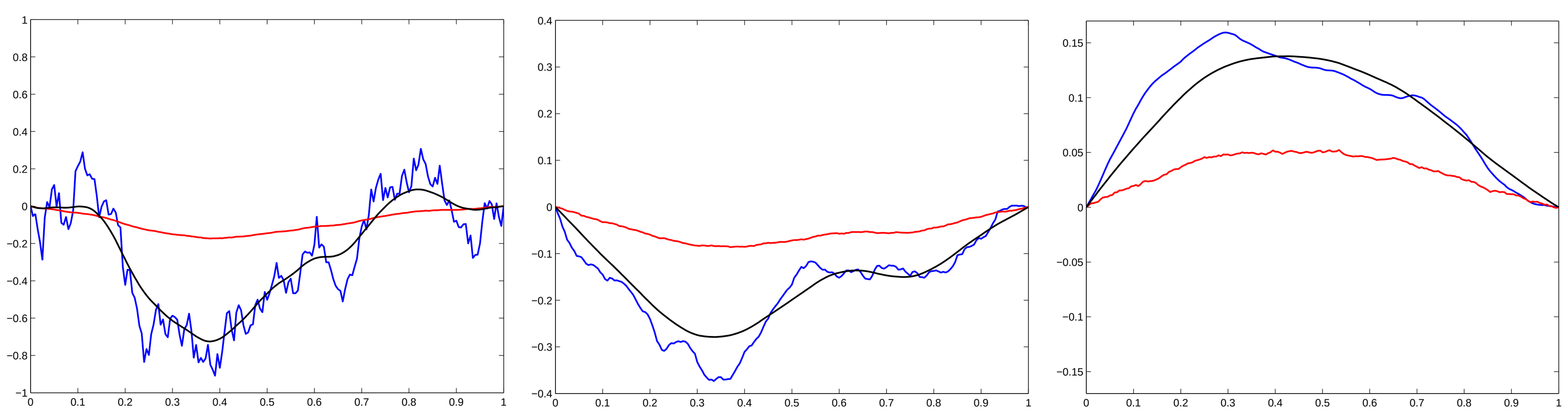
Let $\mathbf{\Delta}$ be the Dirichlet Laplacian in $I = (0, 1)$ and $\{k^2, \phi_k\}$ its orthonormal eigenbasis. Define the "Sobolev-like" spaces

$$\mathcal{H}^s = \{u : \sum_{k=1}^{\infty} k^{2s}\langle u, \phi_k\rangle^2 < \infty\}, \qquad ||u||_s^2 = \sum_{k=1}^{\infty} k^{2s}\langle u, \phi_k\rangle^2.$$

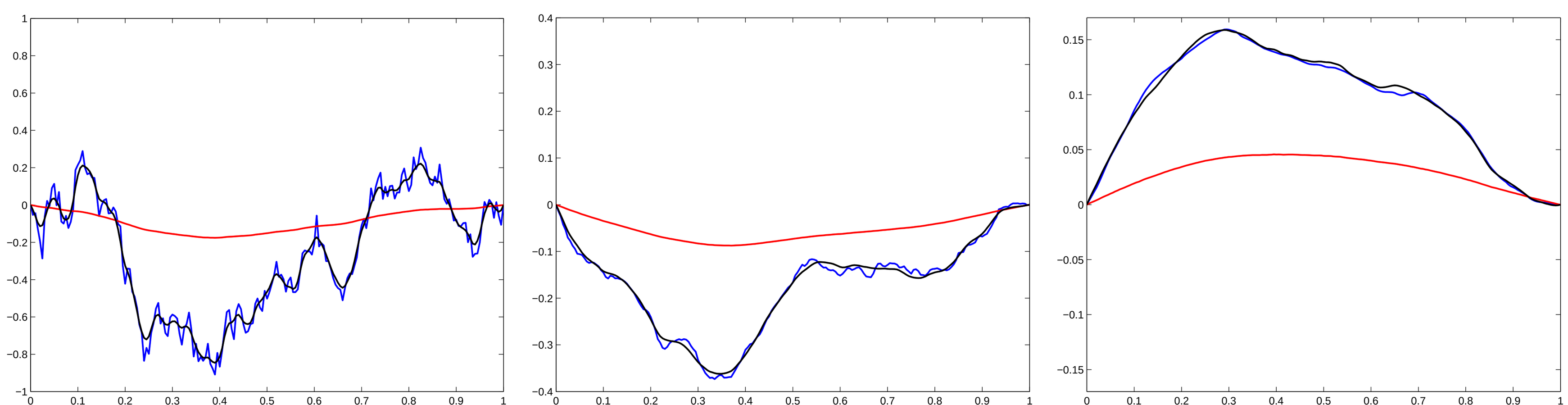
Suppose that in (IP) we have $C_0 = (-\mathbf{\Delta})^{-\frac{1}{2}-\alpha}$, $\mathcal{A}^{-\ell} = C_0^{\frac{2\ell}{1+2\alpha}}$ and $C_1 = C_0^{\frac{2\beta}{1+2\alpha}}$ where $\alpha > 0, \beta \geq 0$, i.e. C_0, $\mathcal{A}^{-\ell}$ and C_1 are diagonalizable in the same eigenbasis.

Following the usual practice in Inverse Problems, we will assume in (DM) that we have the a-priori information on the regularity of $u^\dagger$, that $u^\dagger \in \mathcal{H}^\gamma, \ \gamma \geq 0$.

Numerical Simulation - Diagonal Case



(a) $\gamma = \frac{1}{2}$, $n = 10^4$, $||m - u^\dagger|| \approx 0.29$ (b) $\gamma = 1$, $n = 10^4$, $||m - u^\dagger|| \approx 0.097$ (c) $\gamma = \frac{3}{2}$, $n = 10^4$, $||m - u^\dagger|| \approx 0.046$



(d) $\gamma = \frac{1}{2}$, $n = 10^6$, $||m - u^\dagger|| \approx 0.11$ (e) $\gamma = 1$, $n = 10^6$, $||m - u^\dagger|| \approx 0.022$ (f) $\gamma = \frac{3}{2}$, $n = 10^6$, $||m - u^\dagger|| \approx 0.0073$

Figure: Realizations of the true solution (blue), the posterior mean (black) and the observation (red) for $\alpha = \frac{1}{4}$, $\beta = \frac{1}{2}$, $\ell = \frac{1}{2}$, for three values of γ and for $n = 10^4$ (1st line) and $n = 10^6$ (2nd line).

Observation: We have a better reconstruction when the truth is more regular.

Posterior Consistency - Diagonal Case

Define $\Delta = \frac{1}{4} + \frac{\alpha}{2} + \ell - \frac{\beta}{2}$, and assume that $\Delta > 0$ (the prior is indeed regularizing).

We study optimal rates of convergence for the different values of γ:

-**Theorem:** If the truth is smooth enough, in particular if $\gamma \geq 0$

i) If the noise and the true solution are not too smooth, $\beta - 2\ell \leq \frac{1}{2}$ and $\gamma < 4\Delta$, for

$$\tau_n = n^{\frac{\alpha-\gamma}{4\Delta+2\gamma-2\alpha+\varepsilon}},$$

we have

$$\mathbb{E}\left|\left|m_n - u^\dagger\right|\right|^2 \asymp n^{\frac{-\gamma}{2\Delta+\gamma-\alpha+\varepsilon}}.$$

If the true solution is too smooth, $\gamma \geq 4\Delta$, then the rate does not improve as γ increases, i.e. the rate saturates at

$$n^{\frac{-4\Delta}{6\Delta-\alpha+\varepsilon}}.$$

ii) If the noise is too smooth, $\beta - 2\ell > \frac{1}{2}$, then the resulting rate is arbitrarily close to n^{-1}.

Observation: in (i) for $\alpha > \gamma$ the prior distribution becomes uniform as $n \to \infty$, $\alpha < \gamma$ it contracts to a Dirac, $\alpha = \gamma$ it stays fixed. In the last case the prior regularity matches the regularity of the true solution so rescaling is redundant.

Simulated Rates of Convergence

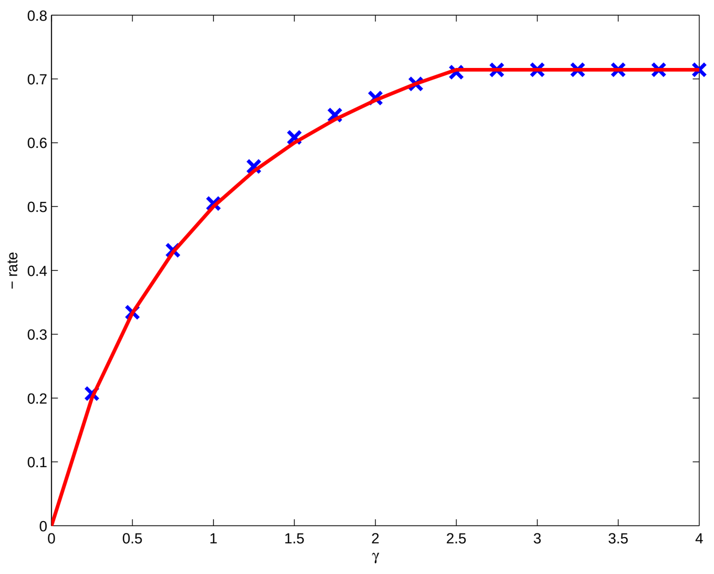


Figure: Rates of convergence of $\mathbb{E}\left|\left|m_n - u^\dagger\right|\right|^2$ plotted against γ. In red we have the rates predicted by the theory and in blue we have estimated rates using the average of 160 realizations.

Observation: The estimated rates saturate at $\gamma = 2.5$ as predicted by the theory.

Posterior Consistency - Non-Diagonal Case

-**Theorem:** Under weaker, non-diagonal assumptions on the equivalence of C_0 to $\mathcal{A}^{-\ell}$ and C_1, using PDE methods and interpolation inequalities we have that for $\theta \in [0, 1]$

$$\mathbb{E}\left|\left|C_0^{\frac{2\ell-\beta-2\theta\Delta}{1+2\alpha}}(m_n - u^\dagger)\right|\right|^2 \leq c\mathbb{E}(\kappa^2)n^{\frac{\theta_2+\theta-2}{\theta_1-\theta_2+2}},$$

where

$$\kappa = \max\left\{\left|\left|C_0^{\frac{2\theta_1\Delta-\beta}{1+2\alpha}}\xi\right|\right|, \left|\left|C_0^{\frac{2(\theta_2-1)\Delta}{1+2\alpha}-\frac{1}{2}}u^\dagger\right|\right|\right\}$$

and where θ_1, θ_2 are chosen so that $\mathbb{E}(\kappa^2) < \infty$.

In the diagonal case the assumptions are trivially satisfied, so we can examine if the general theory agrees with the rates obtained in the diagonal case.

- For $\beta - 2\ell \leq 0$, we can choose $\theta = \frac{2\ell-\beta}{2\Delta}$ in the general method to get the rate for $\mathbb{E}\left|\left|m_n - u^\dagger\right|\right|^2$. Compared to the diagonal case, we require higher regularity of the true solution in order to have convergence ($\gamma \geq \frac{1}{2} + \alpha$) and the rate saturates earlier ($\gamma = \frac{1}{2} + \alpha + 2\Delta$). For $\gamma \in \left[\frac{1}{2} + \alpha, \frac{1}{2} + \alpha + 2\Delta\right]$ the rate agrees with the diagonal case.

The discrepancies can be explained by the fact that in the general method, the choice of θ_1, θ_2 which determines both the minimum requirement on the regularity of $u^\dagger$ and the saturation point, is independent of the choice of θ. This means that on the one hand to get convergence of $\mathbb{E}\left|\left|m_n - u^\dagger\right|\right|^2$ we require conditions which secure the convergence of stronger norms of the error and on the other hand the saturation rate for $\mathbb{E}\left|\left|m_n - u^\dagger\right|\right|^2$ is the same as the saturation rates for weaker norms of the error.

- For $\beta - 2\ell > 0$, the general method provides rates only for stronger norms of the error, hence it secures the convergence of $\mathbb{E}\left|\left|m_n - u^\dagger\right|\right|^2$ but at suboptimal rates.

References

S. Agapiou, S. Larsson, A. M. Stuart,
In preparation.

B. T. Knapik, A. W. van der Vaart, J. H. van Zanten,
Bayesian Inverse Problems, (2011).

P. Diaconis and D. Freedman,
On the consistency of Bayes Estimates, Ann. Statist. 14 (1986), no. 1, 1-67.