Bayesian Inverse Problems with Non-Gaussian Noise

Remo Kretschmann

Universität Duisburg-Essen

e-mail: remo.kretschmann@uni-due.de   web: https://udue.de/kretschmann

Background

Maximum a posteriori (MAP) estimates for Bayesian inverse problems between **infinite-dimensional** Hilbert spaces. Under certain assumptions MAP estimates can be characterised as minimisers of the Onsager–Machlup functional of the posterior distribution, see [1]. In [2], the inverse heat equation in general form was studied in a Bayesian setting with Gaussian noise.

Questions:

- Is there a similar variational characterisation of MAP estimates in case of **non-Gaussian noise**?
- Is there a connection between MAP estimates and ℓ^1-minimisation in case of **Laplacian noise**.
- Inverse heat equation with Laplacian noise: Does the MAP estimator converge in the small noise limit?

Laplacian measures on Hilbert spaces

For an orthonormal basis $(e_k)_{k\in\mathbb{N}}$ of a separable Hilbert space X and $(\sigma_k)_{k\in\mathbb{N}}$ with $\sigma_k \geq 0$ and $\sum_{k=1}^{\infty} \sigma_k^2 < \infty$ define the random variable

$$\xi := \sum_{k=1}^{\infty} \xi_k \sigma_k e_k,$$

where $(\xi_k)_{k\in\mathbb{N}}$ are i.i.d. **standard Laplacian random variables** in $\mathbb{R}$. It has mean 0 and covariance operator $Qx := \sum_{k=1}^{\infty} \sigma_k^2 (x, e_k)_X e_k$. Denote its distribution by $\mathcal{L}_Q$.

For $a \in X$, the shifted measure $\mathcal{L}_{a,Q} := \mathcal{L}_Q(\cdot - a)$ is **absolutely continuous** w.r.t. $\mathcal{L}_Q$ if and only if $a \in Q^{\frac{1}{2}}(X)$, and in this case

$$\frac{\mathrm{d}\mathcal{L}_{a,Q}}{\mathrm{d}\mathcal{L}_Q}(x) = \exp\left(-\sqrt{2} \sum_{k=1}^{\infty} \frac{|(x, e_k)_X - (a, e_k)_X| - |(x, e_k)_X|}{\sigma_k} \right).$$

A severely ill-posed linear problem (P1)

Consider operator equation with additive noise,

$$y = \exp(-A)x + \eta,$$

where A is a densely defined, self-adjoint, surjective linear operator in a separable Hilbert space X with the following properties:

- The eigenvalue problem $A\varphi = \alpha\varphi$ has a countable set of solutions $(\varphi_k)_{k\in\mathbb{N}}$ that forms an **orthonormal basis** of X.
- For some $d \in \mathbb{N}$ and $C^-, C^+ > 0$ the associated eigenvalues $(\alpha_k)_{k\in\mathbb{N}}$ satisfy the **growth condition**

$$C^- k^{\frac{2}{d}} \leq \alpha_k \leq C^+ k^{\frac{2}{d}} \quad \text{for all } k \in \mathbb{N}.$$

Model prior and noise as follows:

- **Centred Gaussian prior** $x \sim \mu_0 := \mathcal{N}_{r^2 A^{-\tau}}$ in X with $r > 0$ and $\tau > \frac{d}{2}$.
- **Centred Laplacian noise** $\eta \sim \mathcal{L}_{b^2 A^{-\beta}}$ in $Y := X$ with $b > 0$ and $\beta > \frac{d}{2}$. Here $\sigma_k := b\alpha_k^{-\beta/2}$ and $e_k := \varphi_k$.

Example: For appropriate bounded open sets $D \subset \mathbb{R}^d$ with Lipschitz boundary, the **inverse heat equation** on D can be expressed in this way by choosing $A = -\Delta$ in $X = L^2(D)$ with domain $\mathcal{D}(A) = H^2(D) \cap H_0^1(D)$. Here Δ denotes the Laplace operator.

The conditional distribution of x given y is called **posterior distribution** and denoted by μ^y. Bayesian inversion yields its density w.r.t. the Gaussian prior distribution μ_0,

$$\frac{\mathrm{d}\mu^y}{\mathrm{d}\mu_0}(x) = \frac{1}{Z(y)} \exp(-\Phi(x; y)),$$

where

$$\Phi(x; y) := \sqrt{2} \sum_{k=1}^{\infty} \frac{|(y, \varphi_k)_X - e^{-\alpha_k}(x, \varphi_k)_X| - |(y, \varphi_k)_X|}{b\alpha_k^{-\beta/2}}.$$

This functional Φ is **Lipschitz continuous** in x.

Variational characterisation of MAP estimates

Consider general Bayesian inverse problem between separable Hilbert spaces X and Y: Given noisy data $y \in Y$, find the unknown $x \in X$. Assumptions:

- **Centred Gaussian prior distribution** μ_0 on X, Cameron–Martin space $(E, \|\cdot\|_E)$.
- The **posterior distribution** μ^y on X has a density w.r.t. μ_0,

$$\frac{\mathrm{d}\mu^y}{\mathrm{d}\mu_0}(x) = \frac{1}{Z(y)} \exp(-\Phi(x; y)).$$

- The functional $x \mapsto \Phi(x; y)$ is **globally Lipschitz continuous**.

Let $B_\varepsilon(x)$ denote the open ball around x with radius ε. A point $\hat{x} \in X$ is called **maximum a posteriori (MAP) estimate** for μ^y, if

$$\lim_{\varepsilon\to 0} \frac{\mu^y(B_\varepsilon(\hat{x}))}{\sup_{x\in X} \mu^y(B_\varepsilon(x))} = 1.$$

A point $\hat{x} \in X$ is a **MAP estimate** for μ^y if and only if $\hat{x} \in E$ and $\hat{x}$ is a **minimiser** of the Onsager–Machlup functional $I \colon E \to \mathbb{R}$,

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

Convergence rate of the MAP estimator in (P1)

Determine MAP estimate for μ^y as unique minimiser of $I(x) = \Phi(x; y) + \frac{1}{2r^2}\|A^{\frac{\tau}{2}} x\|_X^2$ in $E = A^{-\frac{\tau}{2}}(X)$. This yields MAP estimator $\hat{x}_{\mathrm{MAP}}$ in explicit form,

$$\hat{x}_{\mathrm{MAP}}(y) = \sum_{k=1}^{\infty} \max\left\{ -\frac{r^2}{b} R_k, \min\left\{ e^{\alpha_k}(y, \varphi_k)_X, \frac{r^2}{b} R_k \right\}\right\} \varphi_k, \tag{1}$$

where $R_k := \sqrt{2}\alpha_k^{\frac{\beta}{2}-\tau} e^{-\alpha_k}$.

Assume that **true solution** $x^\dagger$ exists and **source condition** $x^\dagger = A^{\frac{\beta}{2}-\tau} e^{-A} w$ holds, where $w \in X$ and $\sup_{k\in\mathbb{N}} |(w, \varphi_k)_X| \leq \rho$. **Choose** $r > 0$ according to $2^{-\frac{1}{2}}\rho b \leq r^2 \leq Cb$ for some $C > 0$. Then the following estimate for the **mean squared error** holds:

$$\mathbb{E}\left[\|\hat{x}_{\mathrm{MAP}}(y) - x^\dagger\|_X^2 \right] \leq 2C\,(\mathrm{Tr} A^{-\tau})\, b.$$

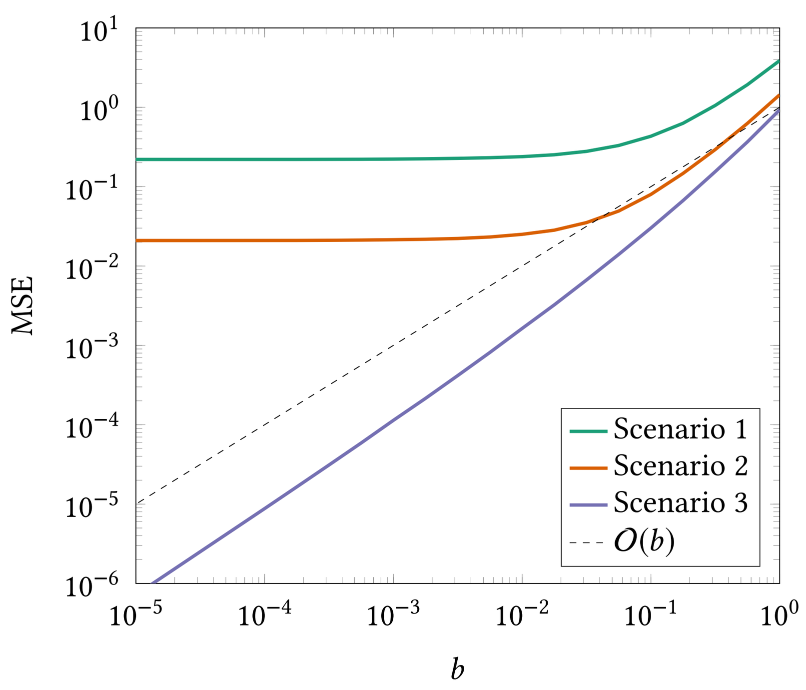
Numerical study of the MAP estimator

The **empirical mean squared error** of the MAP estimator (1) for the 1D inverse heat equation on $(0, 1)$ with $A = 0.002\Delta$, $\tau = 0.55$, $\beta = 0.65$ and $r \propto b^{\frac{1}{2}}$, based on 10000 noise samples. Let w be a piecewise constant function.

Scenario 1: $x^\dagger = w$,

Scenario 2: $x^\dagger = C_2 A^{-\frac{\tau}{2}} w$,

Scenario 3: $x^\dagger = C_3 A^{\frac{\beta}{2}-\tau} e^{-A} w$.



References

[1] M. Dashti, K. J. Law, A. M. Stuart, and J. Voss. MAP estimators and their consistency in Bayesian nonparametric inverse problems. *Inverse Problems*, 29(9):095017, 2013.

[2] M. Dashti and A. M. Stuart. The Bayesian approach to inverse problems. In R. Ghanem, D. Higdon, and H. Owhadi, editors, *Handbook of Uncertainty Quantification*, pages 311–428. Springer International Publishing, Cham, 2017.