# Generalised Modes in Bayesian Inverse Problems

Christian Clason[1] Tapio Helin[2] **Remo Kretschmann**[1]
Petteri Piiroinen[2]

[1]Faculty of Mathematics, University Duisburg-Essen

[2]Department of Mathematics and Statistics, University of Helsinki

Chemnitz Symposium on Inverse Problems 2018
Chemnitz, 27 Sept 2018

# Outline


1 Introduction

2 Generalised modes

3 Variational characterisation of generalised MAP estimates

# Problem setting

UNIVERSITÄT DUISBURG ESSEN Open-Minded

**Bayesian statistical inverse problem**

Assumptions:

- Probability distribution $\mu_0$ of **unknown** $x$ (**prior distribution**)
- Conditional probability distribution of **data** $y$ given unknown $x$

Given:

- Observed data $y$

Find:

- Conditional probability distribution $\mu^y$ of $x|y$ (**posterior distribution**)

**MAP estimate**: Use **mode** of posterior distribution $\mu^y$ as **point estimate** for posterior $x|y$.

# Modes in finite-dimensional setting

Continuous random variables:

- Posterior $x|y$ has **density w.r.t. Lebesgue measure**.
- Modes of posterior distribution $\mu^y$ given by **maximisers of posterior density**, or, equivalently, by minimisers of negative logarithmic posterior density.

**Problem:** There is no Lebesgue measure on **infinite-dimensional** separable Banach spaces.

# Modes in infinite-dimensional setting

Let $\mu$ be a probability measure on a **separable Banach space** $X$ and let $B^\delta(x) \subset X$ denote the open ball centred at $x \in X$ with radius $\delta$.

**Definition**

A point $\hat{x} \in X$ is called a **(strong) mode** of $\mu$ if

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

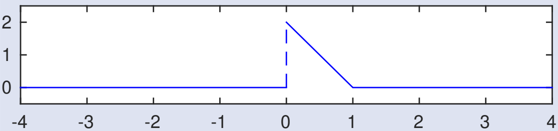
**[Dashti et al 2013]:** In case of **Gaussian prior** and under certain conditions, modes of posterior distribution $\mu^y$ are precisely **minimisers** of a **suitable objective functional**.

# Problems with this definition

## Example

Consider probability measure $\mu$ on $\mathbb{R}$ with Lebesgue density

$$p(x) = \begin{cases} 2(1-x), & \text{if } x \in [0,1], \\ 0, & \text{otherwise.} \end{cases}$$




Then $\mu$ does not have a mode in 0, because

$$\lim_{\delta \to 0} \frac{\mu(B^\delta(0))}{\sup_{x \in X} \mu(B^\delta(x))} = \frac{1}{2}.$$

# Strict bounds

- There are applications where **strict bounds** on the admissible values of the unknown $x$ emerge in a natural way, e.g., X-ray imaging, electrical impedance tomography.
- **Question:** How to generalise the notion of a mode in such a way that it covers what we would intuitively consider a mode in those cases?

# Outline

1 Introduction

2 Generalised modes

3 Variational characterisation of generalised MAP estimates

# Definition of generalised mode

Let $\mu$ be a probability measure on a separable Banach space $X$.

**Definition**

A point $\hat{x} \in X$ is called a **generalised mode** of $\mu$ if for every sequence $\{\delta_n\}_{n\in\mathbb{N}} \subset (0,\infty)$ with $\delta_n \to 0$ there exists a **qualifying sequence** $\{w_n\}_{n\in\mathbb{N}} \subset X$ with $w_n \to \hat{x}$ in $X$ and

$$\lim_{n\to\infty} \frac{\mu(B^{\delta_n}(w_n))}{\sup_{x\in X} \mu(B^{\delta_n}(x))} = 1.$$

- Every strong mode is a generalised mode.
- In the previous example 0 is a generalised mode with $w_n := \delta_n$.
- For Gaussian measures the strong mode is the only generalised mode.

# Relation between modes and g-modes

Criteria for coincidence of strong and generalised modes:

1. Convergence rate of the qualifying sequence $w_n$ and equicontinuity condition.
2. Convergence of qualifying sequence $w_n$ in subspace topology and equicontinuity condition.

# Outline


1 Introduction

2 Generalised modes

3 Variational characterisation of generalised MAP estimates

# Uniform prior measure

UNIVERSITÄT DUISBURG ESSEN Open-Minded

Set

$$X := c_0 := \{\{x_k\}_{k\in\mathbb{N}} \in \ell^\infty : \lim_{k\to\infty} x_k = 0\}$$

and $\|\cdot\|_X := \|\cdot\|_\infty$. Define the $X$-valued random variable

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

where

- $\{\xi_k\}_{k\in\mathbb{N}}$ are i.i.d. **uniformly distributed** on $[-1, 1]$,
- $\gamma_k \geqslant 0$ for all $k \in \mathbb{N}$ and $\gamma_k \to 0$,
- $\{e_k\}_j = 1$ for $j = k$ and 0 otherwise.

Now define probability measure $\mu_\gamma$ on $(X, \mathcal{B}(X))$ by

$$\mu_\gamma(A) := \mathbb{P}[\xi \in A] \quad \text{for all } A \in \mathcal{B}(X).$$

**Assumptions:**

- **Uniform prior distribution** $\mu_0 := \mu_\gamma$ on $X := c_0$ as defined before for non-negative weights $\{\gamma_k\}_{k\in\mathbb{N}}$ with $\gamma_k \to 0$.
- For given data $y$ in separable Hilbert space $Y$, the **posterior distribution** $\mu^y$ is absolutely continuous w.r.t. $\mu_0$ and

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

  for all $A \in \mathcal{B}(X)$, where $Z(y) := \int_X \exp(-\Phi(x; y))\mu_0(\mathrm{d}x)$ and $\Phi\colon X \times Y \to \mathbb{R}$.
- $0 < Z(y) < \infty$.
- The **negative log-likelihood** $\Phi(\cdot; y)$ is Lipschitz continuous on bounded sets.

# Generalised MAP estimates

Set $E_\gamma := \{x \in X : |x_k| \leqslant \gamma_k \text{ for all } k \in \mathbb{N}\}$ and define the functional $I^y \colon X \to \mathbb{R} \cup \{\infty\}$,

$$I^y(x) := \Phi(x; y) + \iota_{E_\gamma}(x),$$

where

$$\iota_{E_\gamma}(x) = \begin{cases} 0, & \text{if } x \in E_\gamma, \\ \infty, & \text{otherwise.} \end{cases}$$

**Theorem**

A point $\hat{x} \in X$ is a **generalised mode** of $\mu^y$ if and only if it is a **minimiser** of $I^y$.

Generalised MAP estimation corresponds to **Ivanov regularisation** with the compact set $E_\gamma$.

# Idea of proof

UNIVERSITÄT DUISBURG ESSEN

Open-Minded

1. For every $\delta > 0$ let $x^\delta$ be a **maximiser** of $\mu^y(B^\delta(\cdot))$ in $X$.
2. The family $\{x^\delta\}_{\delta>0}$ contains a **convergent subsequence** and the **limit** of every convergent subsequence **minimises** $I^y$.
3. Every **minimiser** of $I^y$ is a **generalised mode** of $\mu^y$ and vice versa.

# Conclusion

- We have conditions for the **coincidence** of strong modes and generalised modes.
- In case of a **uniform prior**, generalised MAP estimates can be characterised as the **minimisers** of a **suitable objective functional**.

**References:**


C. Clason, T. Helin, R. Kretschmann, and P. Piiroinen.
Generalized modes in Bayesian inverse problems.
Submitted, 2018. arXiv:1806.00519.

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.