Besov Priors for Bayesian Inverse problems

Masoumeh Dashti

Mathematics Department
University of Sussex

December 13, 2011, Isaac Newton Institute

Collaboration with
Stephen Harris (Edinburgh) and Andrew Stuart (Warwick)
Outline

1. Inverse problems for functions, Bayesian regularization
2. Besov prior measures, and wellposedness of posterior
3. Approximate posterior: convergence properties; weak error estimates
4. Example: Bayesian regularization of an elliptic problem
Outline

1. Inverse problems for functions, Bayesian regularization
2. Besov prior measures, and wellposedness of posterior
3. Approximate posterior: convergence properties; weak error estimates
4. Example: Bayesian regularization of an elliptic problem
Inverse problems and ill-posedness

- Consider

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

\( \mathcal{G} : X \rightarrow Y, \ X, Y \) Banach spaces
\( \eta : Y \)–valued random variable

Given \( y, \mathcal{G} \) and statistical properties of \( \eta \),
Find an estimate of \( u \)

---

- Example: Consider

\[ \nabla \cdot (e^{u(x)} \nabla p(x)) = 0, \quad x \in D \subset \mathbb{R}^d, \quad d \leq 3, \]
\[ p = \phi \quad \text{on} \ \partial D \]

Find \( u \)

Given noisy observations of \( p \),
\[ y_j = p(x_j) + \eta_j, \quad \text{so} \quad \mathcal{G} : X \rightarrow \mathbb{R}^K \]
\[ u \mapsto \{p(x_j)\}_{j=1}^K \]
Classical regularization

- Let $y = G(u)$ given $y \in Y$ and for $u \in X$ be ill-posed.
- Best approximate solution:
  $$\arg\min_{u \in X} \| y - G(u) \|_Y$$
- Regularized solution:
  $$\arg\min_{u \in X} \| y - G(u) \|_Y^2 + \alpha J(u),$$
  e.g. $J(u) = \| u \|_E^2$ for $E \subset X$

$\{ y, G, \text{ appropriate prior information} \} \rightarrow$
reasonable approximation of ‘true’ solution
Bayesian regularization

- Find distribution of \( u \in X \) given \( y \in Y \)

\[
y = g(u) + \eta, \quad \eta \sim \mathcal{N}(0, \Gamma)
\]

- \( \mathbb{P}(y|u) \) known,
- Prior \( \mu_0(du) \) on \( u \) with \( \mu_0(X) = 1 \)
- Solution: Posterior \( \mu^y(du) \) on \( u \):

\[
\frac{d\mu^y}{d\mu_0}(u) \propto \mathbb{P}(y|u) \propto \exp(-\Phi(u, y)).
\]

\{y, statistical properties of \( \eta, g, \mu_0 \}\n\rightarrow \mu^y
Outline

1. Inverse problems for functions, Bayesian regularization
2. Besov prior measures, and wellposedness of posterior
3. Approximate posterior: convergence properties; weak error estimates
4. Example: Bayesian regularization of an elliptic problem
Prior measure

Besov prior:

- based on wavelet expansions
- appropriate when we expect $u$ to be smooth with a few local irregularities
- promotes sparsity
Prior measure

We define the prior measure through Karhunen-Loève expansion of its draws:

Besov priors (Lassas, Saksmas, Siltanen 2009)

\[ u(x) = \sum_{j=1}^{\infty} \left( \frac{1}{\kappa} \right)^{\frac{1}{q}} \lambda_j^{\frac{1}{q}} \xi_j \psi_j(x), \quad \lambda_j = j^{-\left(\frac{q s}{d} + \frac{q}{2} - 1\right)} \]

\( x \in \mathbb{T}^d, \, q \in [1, \infty), \, s > 0, \, \kappa > 0 \)

- \( \{\xi_j\}_{j \in \mathbb{N}} \): i.i.d, and \( \xi_j \sim c_q \exp(-\frac{1}{2}|x|^q) \)
- \( \{\psi_j\}_{j \in \mathbb{N}} \): \( r \)-regular \((r > s)\) wavelet basis of \( L^2(\mathbb{T}^d) \) for \( q \geq 1 \) allowing choice of Fourier basis for \( q = 2 \)
Prior measure

\[ B_{qq}^s = \{ u \in L^q(\mathbb{T}^d) : \sum_{j=1}^{\infty} j^{\frac{qs}{d} + \frac{q}{2} - 1} |(u, \psi_j)|^q < \infty \} \]

\[ \|u\|_{B_{qq}^s} \]

\[ \mu_0 \text{ is a } (\kappa, B_{qq}^s) \text{ measure if draws } u \text{ from } \mu_0 \text{ satisfy} \]

\[ u(x) = \sum_{j=1}^{\infty} \left( \frac{1}{\kappa} \right)^{\frac{1}{q}} \lambda_j^{\frac{1}{q}} \xi_j \psi_j(x), \quad \lambda_j = j^{-\left( \frac{qs}{d} + \frac{q}{2} - 1 \right)} \]

Let \( s > d/q \), then for any \( t < s - d/q \)

- \( \mu_0(B_{qq}^t) = 1 \)
- \( \mu_0(C^t) = 1 \)
• What is an upper bound for $\mu_0$-integrable functions?

$\int_X \exp(\alpha \|u\|_{B_{qq}^t}) \mu_0(du) < \infty$, (Lassas et al 2009)

• How about $\int_X \exp(\alpha \|u\|_X)$ for any $X$ with $\mu_0(X) = 1$?

If $s > \frac{d}{q}$, then for any $t < s - \frac{d}{q}$

$\int_X \exp(\lambda \|u\|_{C^t}) \mu_0(du) < \infty$, for any $\lambda < \kappa c(t, q, d)$. 
• $\int \exp(\lambda \|u\|_{C^t})\mu_0(du) < \infty$:
  - In an $r$-regular wavelet basis
    \[
    \|u\|_{C^t} = \sup_{j \in \mathbb{N}} j^{(t-s)/d} |\xi_l|
    \]
  - An appropriate change of proof of J.P. Kahane 1968 for Rademacher series gives the result.

• As a result
  \[
  \int \exp(\lambda \|u\|_{C^t})\mu_0(du) < \infty \quad \text{for any} \quad \theta \in [1, \infty].
  \]
Observation operator

Conditions on the Potential $\Phi$:

- there exists a $c > 0$ and $\forall r > 0$ a $M = M(r) > 0$ such that,

\[
\Phi(u; y) \geq -c\|u\|_X + M \quad \forall u \in X, \|y\|_Y < r
\]
\[
\Phi(u; y) \leq K(r) \quad \text{for } \max\{\|u\|_X, \|y\|_Y\} < r
\]

- for every $r > 0$ there is a $K = K(r) > 0$ such that,

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

- there exists $c > 0$ and for every $r > 0$ a $K = K(r) > 0$ s.t.

\[
|\Phi(u; y_1) - \Phi(u; y_2)| \leq Ke^{c\|u\|_X} \|y_1 - y_2\|_Y \quad \forall u \in X, \|y_i\|_Y < r
\]
Observation operator

Conditions on the Potential $\Phi$:

- there exists a $c > 0$ and $\forall r > 0$ a $M = M(r) > 0$ such that,
  \[
  \Phi(u; y) \geq -c\|u\|_X + M \quad \forall u \in X, \|y\|_Y < r \\
  \Phi(u; y) \leq K(r) \quad \text{for } \max\{\|u\|_X, \|y\|_Y\} < r
  \]

- for every $r > 0$ there is a $K = K(r) > 0$ such that,
  \[
  |\Phi(u_1; y) - \Phi(u_2; y)| \leq K \|u_1 - u_2\|_X \quad \forall \|u_i\|_X, \|y\|_Y < r
  \]

- there exists $c > 0$ and for every $r > 0$ a $K = K(r) > 0$ s.t.
  \[
  |\Phi(u; y_1) - \Phi(u; y_2)| \leq Ke^c\|u\|_X \|y_1 - y_2\|_Y \quad \forall u \in X, \|y_i\|_Y < r
  \]
Wellposedness of the posterior

Theorem

Let
- $\Phi$ satisfy POTENTIAL CONDITIONS,
- $\mu_0$ be a Besov or Gaussian measure with $\mu_0(X) = 1$.

Then

$\mu^y$ is well-defined and

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

where $C = C(r)$ with $\max\{|y|, |y'|\} \leq r$.

Metric $d_{\text{Hell}}$ is a useful one because:

$$\|E^\mu f - E^\nu f\| \leq 2\left( \|E^\mu f\|^2 + \|E^\nu f\|^2 \right)^{\frac{1}{2}} d_{\text{Hell}}(\mu, \nu).$$
Outline

1. Inverse problems for functions, Bayesian regularization
2. Besov prior measures, and wellposedness of posterior
3. Approximate posterior: convergence properties; weak error estimates
4. Example: Bayesian regularization of an elliptic problem
Approximation of the posterior

Consider $\mu$ and $\mu^N$:

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

**Theorem**

Let

- $\Phi$ and $\Phi^N$ satisfy POTENTIAL CONDITIONS uniformly in $N$,
- $\mu_0$ be a Besov or Gaussian measure with $\mu_0(X) = 1$,
- and assume also that $\exists K, c > 0$ s.t

  $$|\Phi(u) - \Phi^N(u)| \leq K \exp(c\|u\|_X) \psi(N)$$

  with $\psi(N) \to 0$ as $N \to \infty$

Then there exists a constant $C$ independent of $N$ such that

$$d_{\text{Hell}}(\mu, \mu^N) \leq C \psi(N).$$
Finite-d approximation of the posterior

- Recall: \( \{\psi_j\} \) is the basis of \( L^2 \) using which \( \mu_0 \) is constructed.
- If \( \Phi^N(\cdot) = \Phi(P^N \cdot) \) with
  \( P^N \) orthogonal proj of \( L^2 \) onto \( W^N = \text{span}\{\psi_1, \ldots, \psi_N\} \)
- Then
  \[
  \frac{d \nu^N}{d \mu_0^N}(u) \propto \exp(-\Phi^N(u)).
  \]
  with \( \nu^N \) and \( \mu_0^N \) living on \( W^N \).

under reasonable conditions on \( \Phi \) and \( \Phi^N \), the weak error:

\[
\| \mathbb{E}^{\mu^Y} F(u) - \mathbb{E}^{\nu^N} F(P^N u) \|_S \to 0
\]

where \( F : X \to S \), \( S \) Banach space.
Outline

1. Inverse problems for functions, Bayesian regularization
2. Besov prior measures, and wellposedness of posterior
3. Approximate posterior: convergence properties; weak error estimates
4. Example: Bayesian regularization of an elliptic problem
Application: elliptic inverse problem

- Consider

\[ \nabla \cdot (e^{u(x)} \nabla p(x)) = 0, \quad x \in T^d, \quad d \leq 3, \]

Assuming \( u \in L^\infty(T^d) \).

- Find \( u \)

Given noisy point-wise measurements of \( p \)

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

\[ \eta_k \sim \mathcal{N}(0, \Gamma), \quad \text{i.i.d} \]

Observation operator \( G : X \rightarrow \mathbb{R}^K \)

\[ u \mapsto \{p(x_k)\}_{k=1}^K \]

and \( \Phi(u, y) = \frac{1}{2} \left| \Gamma^{-\frac{1}{2}} (y - G(u)) \right|^2 \)
Elliptic problem – Wellposedness

Let $X = C^t(D), \, t > 0$.

Then $\Phi$ satisfies **POTENTIAL CONDITIONS**.

Choose $\mu_0$ a $(\kappa, B^s_{qq})$ measure with

\[ q \geq 1, \, s > t + d/q \]

and $\kappa > 4c_0/c(t, q, d)$

so that $\mu_0(C^t) = 1$.

- Then $\mu^y$ is absolutely continuous with respect to $\mu_0$ with
  \[
  \frac{d\mu^y}{d\mu_0}(u) \propto e^{-\frac{1}{2} \left| \Gamma^{-\frac{1}{2}} (y - g(u)) \right|^2}, \quad \text{and} \quad d_{\text{Hell}}(\mu^y, \mu'^y) \leq C |y - y'|.
  \]

- Weak error for $N$-dim posterior is of order $N^{-t/d}$.
Summary

- Studied well-posed Bayesian formulation for inverse problems, when the prior is a Besov measure,
  
  *motivation*: sparsity promoting features of wavelet bases
  
  *key technical tool*: Fernique-like result for Besov measures

- Ongoing: Relationship between the MAP estimator and the minimizer of

\[
\frac{1}{2} \| y - G(u) \|^2_Y + \frac{1}{2} \| u \|_{B^{s}_{qq}}^q .
\]


• M. Dashti, H. Harris, A. M. Stuart, Besov priors for Bayesian inverse problems, Submitted.