A Metropolis-Hastings algorithm for posterior measures with self-decomposable priors

Bamdad Hosseini
Computing+Mathematical Sciences
CalTech
www.bamdadhosseini.org

LMS & CRISM
Summer School on Computational Statistics

July 10, 2018

Goal

$$\frac{\mathsf{d}\nu}{\mathsf{d}\mu}(u) = \frac{1}{Z}\exp(-\Psi(u))$$

- Probability measures ν, μ on Hilbert space X.
- Bayesian inverse problems.
- Algorithms exist when μ is Gaussian [1].
- Large class of consistent approximations [2].
- Function space algorithms seem to scale better.

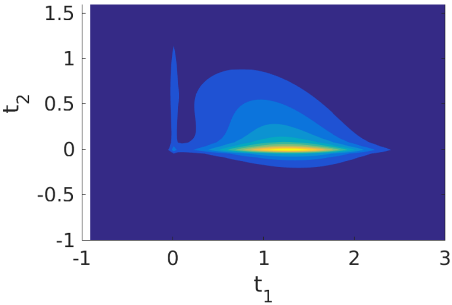
---

[1] S. L. Cotter et al. "MCMC methods for functions: modifying old algorithms to make them faster". In: *Statistical Science* 28.3 (2013), pp. 424–446.
[2] B. Hosseini and J. E. Johndrow. "Analysis of a function space MCMC algorithm for posteriors with self-decomposable priors". In preparation.

Challenges

$$\frac{\mathrm{d}\nu}{\mathrm{d}\mu}(u) = \frac{1}{Z}\exp(-\Psi(u))$$

- Infinite (high) dimensional X.
- Complex negative log-likelihood Ψ.
- Non-Gaussian μ.



Idea

Metropolis-Hastings

1. Start with $u^{(0)} \in X$.
2. Propose $v^{(j+1)} \sim Q(u^{(j)}, \mathrm{d}v)$.
3. Set $u^{(j+1)} = v^{(j+1)}$ with probability $a(u^{(j)}, v^{(j+1)})$.
4. Otherwise reject.
5. Set $j \leftarrow j + 1$ and go to step 2.

- Prior preserving proposal

$$Q\mu = \mu.$$

- Acceptance ratio

$$a(u, v) = \min\{1, \exp(\Psi(u) - \Psi(v))\}.$$

- Reversible on X [3].
- **Can we find such Q? and is it tractable?**

---

[3] L. Tierney. "A note on Metropolis-Hastings kernels for general state spaces". In: *Annals of Applied Probability* 8.1 (1998), pp. 1–9.

ARSD algorithm

- Yes, if μ is self-decomposable (SD).
- Limit distributions of AR(1) processes.
- There exists innovation μ_β for $\beta \in (0,1)$ s.t.

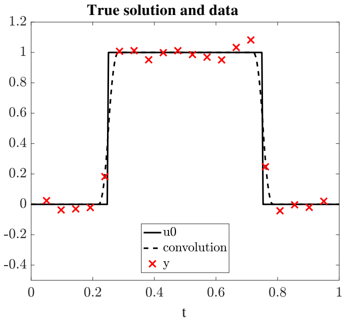
$$\xi \stackrel{d}{=} \beta\tilde{\xi} + \zeta, \qquad \xi, \tilde{\xi} \sim \mu, \quad \zeta \sim \mu_\beta.$$

- Ex: Exponential, Gaussian, Laplace, Gamma, etc.
- AR(autoregressive) SD(self-decomposable).

ARSD algorithm

1. Start with $u^{(0)} \in X$.
2. Propose $v^{(j+1)} = \beta u^{(j)} + \xi, \qquad \xi \sim \mu_\beta$.
3. Set $u^{(j+1)} = v^{(j+1)}$ with probability $\min\{1, \exp(\Psi(u^{(j)}) - \Psi(v^{(j+1)}))\}$.
4. Otherwise reject.
5. Set $j \leftarrow j+1$ and go to step 2.

Deconvolution



- Pointwise observations of $\varphi * u_0$.
- Additive Gaussian noise.
- Data $y \in \mathbb{R}^m$,

$$y = \mathcal{G}(u_0) + \eta, \qquad \eta \sim \mathcal{N}(0, \sigma^2 I).$$

Deconvolution

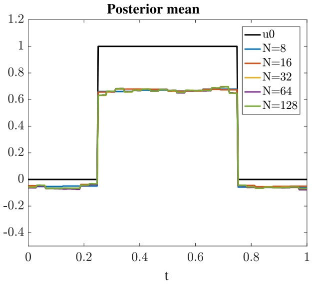
- $X = L^2(\mathbb{T})$.
- $\Psi(u) := \frac{1}{2\sigma^2}\|\mathcal{G}(u) - y\|_2^2, \qquad \sigma > 0.$

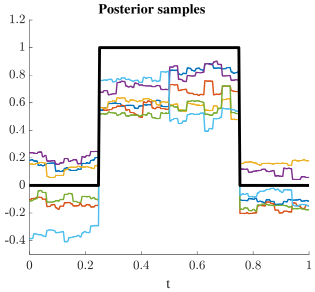
- $\{\psi_k\}_{k=1}^{\infty}$– Haar wavelet system.
- For $j = 1, 2, 3, \cdots$ and $m = 0, 1, 2, \cdots, 2^j - 1$

$$\gamma_0 = \gamma_1 = 1, \qquad \gamma_{2^j+m} = 2^{-2j}.$$

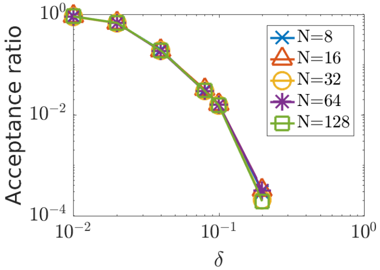
- $\mu = \mathrm{Law}\left[\sum_{k=1}^{N} \gamma_k \eta_k \psi_k\right].$
- $\eta_k = \xi_k - \xi_k', \qquad \xi_k, \xi_k' \sim \mathrm{Gamma}(2/3, 1).$
- Well-defined as $N \to \infty$.

Deconvolution: posterior statistics

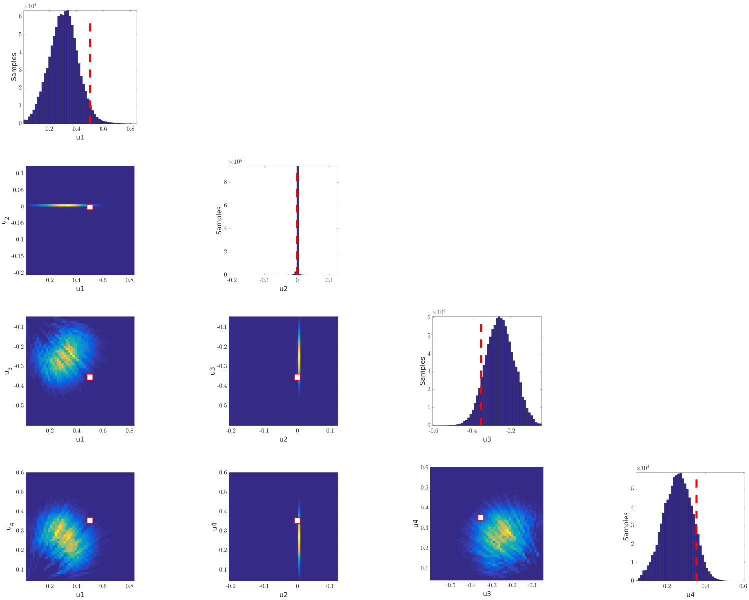




- $\beta = \sqrt{1 - \delta^2}$



Deconvolution: posterior marginals on wavelet modes (u_i, u_j)



Thank you

B. Hosseini. "A Metropolis-Hastings algorithm for posterior measures with self-decomposable priors". In: (2018). arXiv:1804.07833

B. Hosseini and J. E. Johndrow. "Analysis of a function space MCMC algorithm for posteriors with self-decomposable priors". In preparation

References

[1] S. L. Cotter et al. "MCMC methods for functions: modifying old algorithms to make them faster". In: *Statistical Science* 28.3 (2013), pp. 424–446.

[2] B. Hosseini and J. E. Johndrow. "Analysis of a function space MCMC algorithm for posteriors with self-decomposable priors". In preparation.

[3] L. Tierney. "A note on Metropolis-Hastings kernels for general state spaces". In: *Annals of Applied Probability* 8.1 (1998), pp. 1–9.

[4] B. Hosseini. "A Metropolis-Hastings algorithm for posterior measures with self-decomposable priors". In: (2018). arXiv:1804.07833.