Stochastic Simulation in Multimodal Posteriors: UQ in ODEs

Centro de Investigación en Matemáticas, A.C.

Javier Enrique Aguilar | javier.aguilar@cimat.mx, jear2412

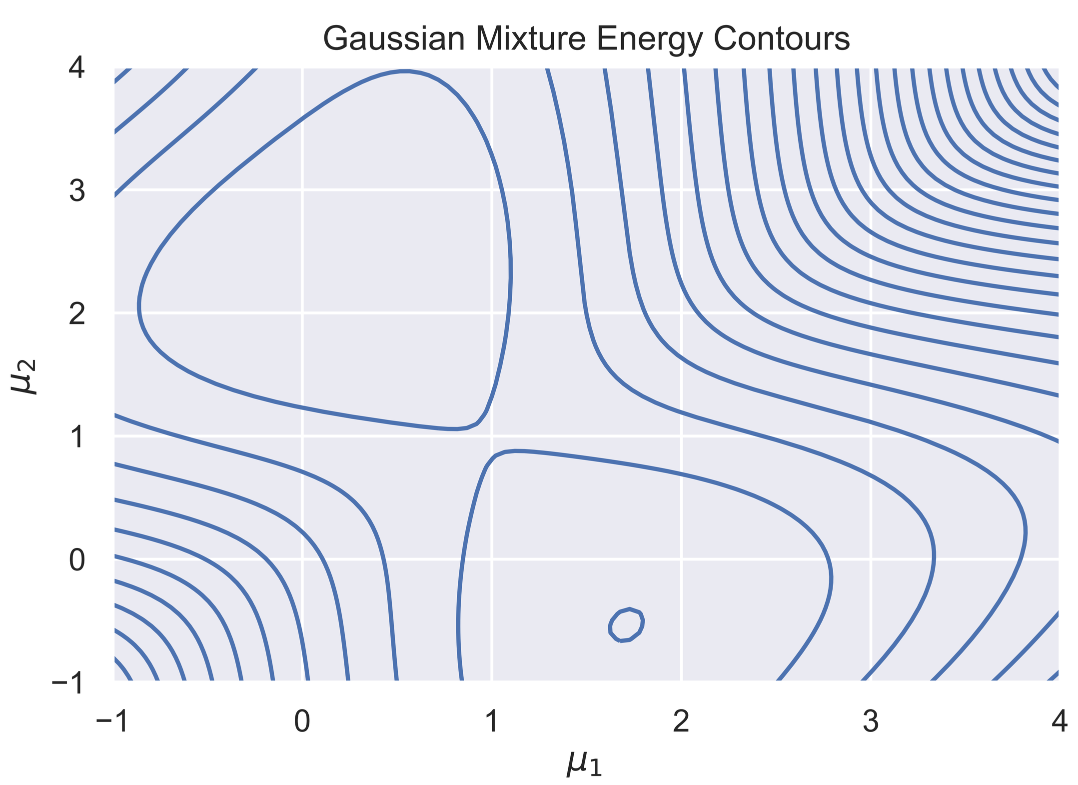
Advisor: José Andrés Christen PhD. | jac@cimat.mx

Introduction

A "basic problem" : Someone asks you to estimate μ_1, μ_2 in the mixture problem

$$X_i|\mu_1,\mu_2 \sim f(\mu_1,\mu_2), \quad \mu_1,\mu_2 \sim U(a,b),$$
$$f(x_i|\mu_1,\mu_2) = p\mathcal{N}(\mu_1,1) + (1-p)\mathcal{N}(\mu_2,1)$$

with $p \neq 0.5$ known.



1. **Obs:** the problem is well specified, but your posterior has a strange shape!
2. **Lesson:** Multimodality can be present even in the simplest problems.
3. **Interest:** Multimodality when solving the inverse problem in ODEs. (Bayesian UQ).

$$y_i = \mathcal{H}(X_\theta(t_i)) + \varepsilon_i, \quad \varepsilon_i \sim_{i.i.d}, i = 1, ..., m$$
$$\frac{dX_\theta}{dt} = F(X_\theta, t, \theta); \quad X_\theta(t_0) = X_0.$$

- $X_\theta(t_i)$ is the **Forward Map**. Complex, non-linear and high dimensional.
- $\mathcal{H} : \mathbb{R}^p \to \mathbb{R}^k$ is the **Observation operator**.
- $\mathcal{H} \circ X_\theta$ induces multimodality!

Metropolis-Hastings

- **MCMC** for the simulation of f, is any simulation method that produces an ergodic Markov chain (X_t) whose stationary distribution is f.
- We give attention to **Metropolis-Hastings** chains. We propose a move through $q(\cdot|\cdot)$ and accept it with probability

$$\rho(x,y) = \left\{1, \frac{f(y)q(x|y)}{f(x)q(y|x)}\right\}.$$

- **Obs:** We use the **Integrated Autocorrelation Time** (IAT) to measure the quality of our chains. (The 'force' of independence our chain has.)
- **SERIOUS Problem:** Chains get stuck. Exploration of entire state space is not possible. NO ergodicity!

Population based MCMC

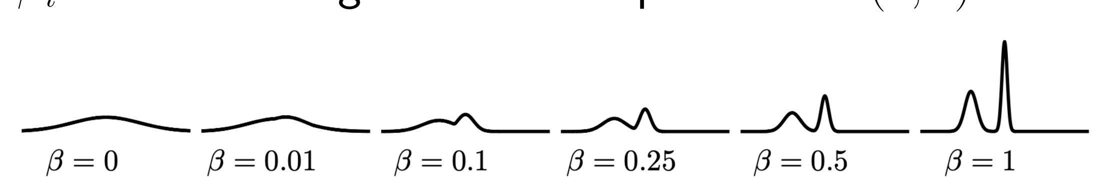
Idea: Extend the state space from $\mathcal{E}$ to $\mathcal{E}^N$

$$f^*(\theta_1, ..., \theta_N|y)d_{\theta_1,...,\theta_N} = \left[\prod_{i=1}^{N} f_i(\theta_i|y)\right] d\theta_1 d\theta_2 ... d\theta_N,$$

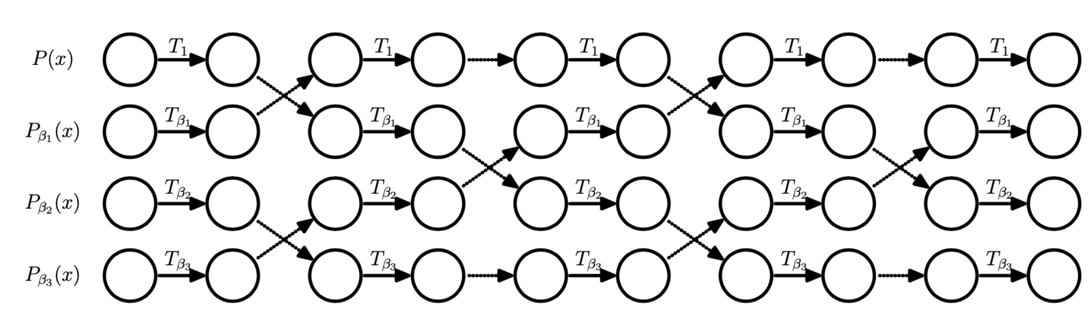
where $f = f_i$ for at least one i.

Parallel Tempering

- **Main Idea:** Use $f_i(\theta|y) \propto f(\theta|y)^{\beta_i}, \beta_i \in B$
- β_i is a smoothing factor or temperature in $(0,1)$

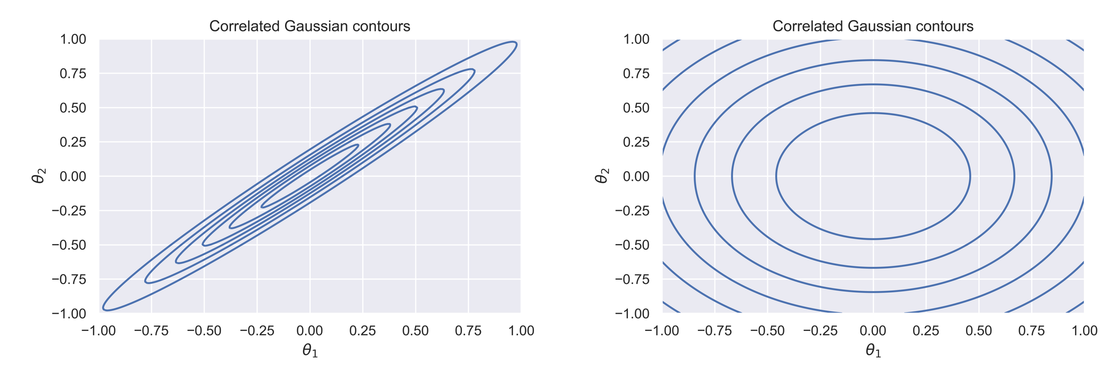


- Heavily based on transferring knowledge from high to low temperatures through an *exchange move*.



Affine Invariant MCMC

- **Main Idea:** be able to sample from densities and affine transformations of them just as equally difficult.
- Important to consider when the correlation structure varies through the state space.
- Some samplers: t-walk, `emcee`



Gradient based MCMC

Main idea: make use of the gradient and Hessian as a means of information about the local geometry of the posterior distribution.

Hamiltonian MC

- Treat $U(\theta) = -\log f(\theta|y)$ as the **potential energy** of a system.
- Introduce auxiliary **momentum** variables p and $K(p)$ to allow Hamiltonian dynamics to operate.
- New target distribution
  $f(\theta, p|y) \propto f(\theta|y)f(p|y) \propto exp\{-H(\theta,p)\}$,
- $H(\theta, p)$ is the **Hamiltonian** and it satisfies $\forall i$

$$\frac{d\theta_i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial \theta_i}.$$

- Simulate *leapfrog dynamics*.
- Highly efficient MCMC with low autocorrelation, but we need the gradient. (Expensive)

Let's see a nice animation `https://chi-feng.github.io/mcmc-demo/app.html`

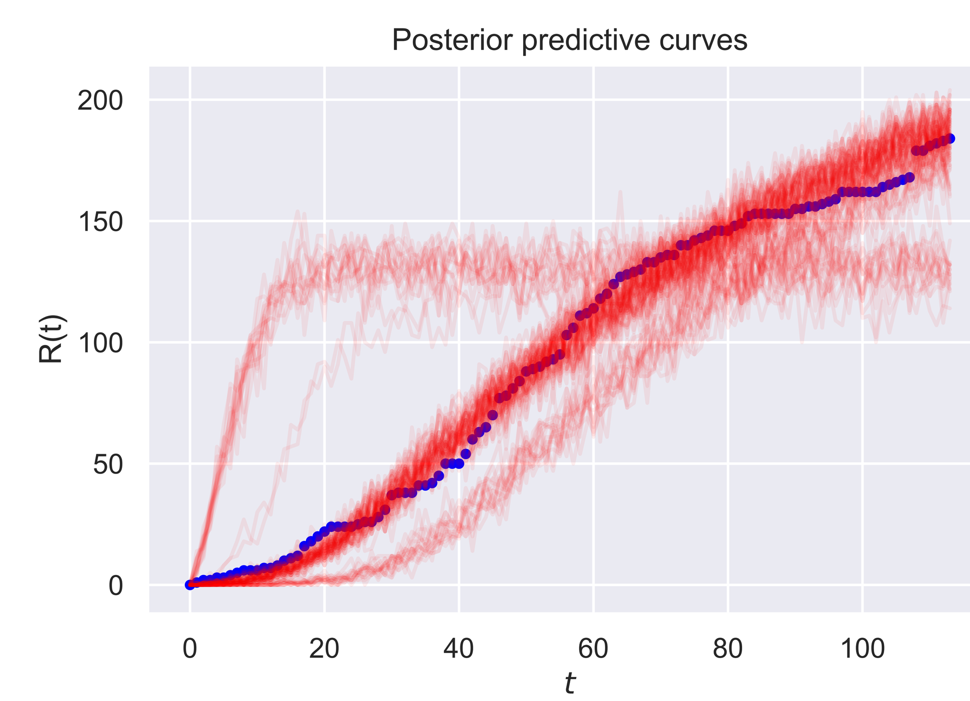
Example

Second Black Plague Eyam, Uk June 19, 1666.
114 days.
SIR Model

$$\frac{dS}{dt} = -\beta S(t)I(t), \quad \frac{dI}{dt} = \beta S(t)I(t) - \alpha I(t), \frac{dR}{dt} = \alpha I(t).$$

Observed y_i is the number of *removed*.

$y_i|\alpha,\beta,I(0) \sim Bin(N, R(t)/N)$, $\alpha,\beta \sim Ga(0,1) I(0) \sim Bin(N, 5/N)$



Conclusions

Simulation from multimodal posteriors can be addressed by combining and implementing several MCMC methods.

References

[1] Ajay Jasra, David A. Stephens, and Christopher C. Holmes. On population-based simulation for static inference. *Statistics and Computing*, 17(3):263–279, 2007.

[2] Christian P Robert, Víctor Elvira, Nick Tawn, and Changye Wu. Accelerating mcmc algorithms. *Wiley interdisciplinary reviews. Computational statistics*, 10(5):e1435–e1435, Sep-Oct 2018.