Markov chain Monte Carlo

PyData London 2024

Calculating Posteriors

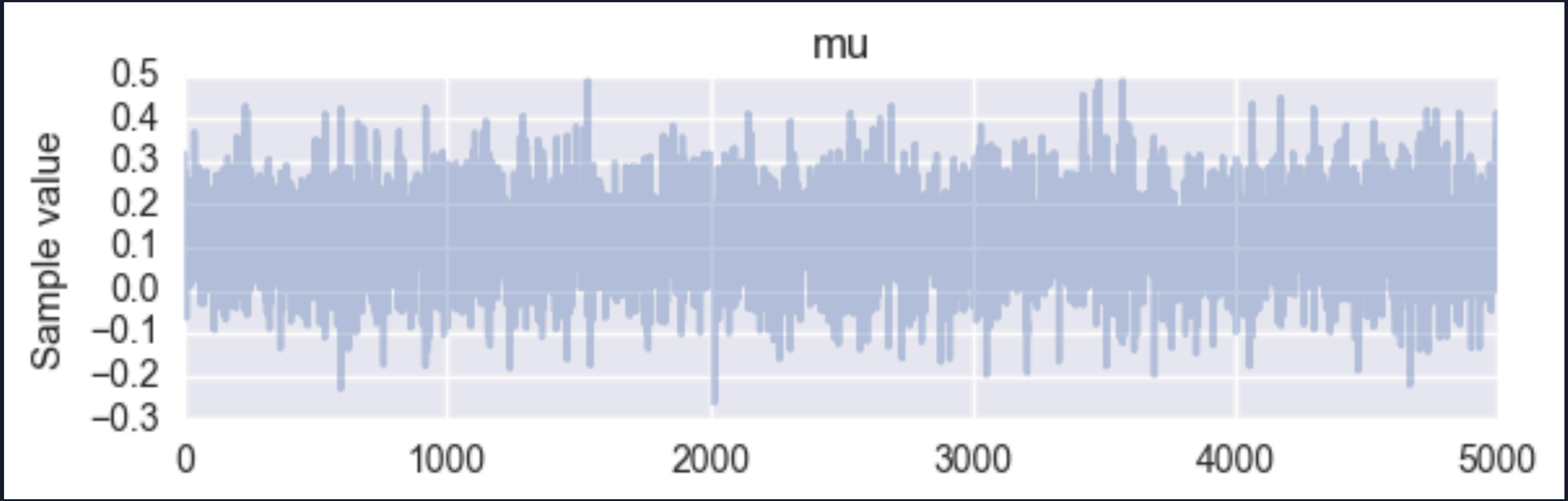
$$Pr(\theta|y) = \frac{Pr(y|\theta)Pr(\theta)}{\int_\theta Pr(y|\theta)Pr(\theta)d\theta}$$

Bayesian approximation

- Maximum *a posteriori* (MAP) estimate
- Laplace (normal) approximation
- Rejection sampling
- Importance sampling
- Sampling importance resampling (SIR)
- Approximate Bayesian Computing (ABC)
- Laplace Approximation

MCMC

Markov chain Monte Carlo simulates a **Markov chain** for which some function of interest is the **unique, invariant, stationary** distribution.



Markov chains

Stochastic process:

$$\{X_t : t \in T\}$$

Markovian condition:

$$Pr(X_{t+1} = x_{t+1} | X_t = x_t, X_{t-1} = x_{t-1}, \ldots, X_0 = x_0)$$

$$= Pr(X_{t+1} = x_{t+1} | X_t = x_t)$$

Reversible Markov chains

When the Markov chain is constructed to satisfy **detailed balance**:

$$\pi(x)Pr(y|x) = \pi(y)Pr(x|y)$$

The π is the limiting distrbution of the chain.

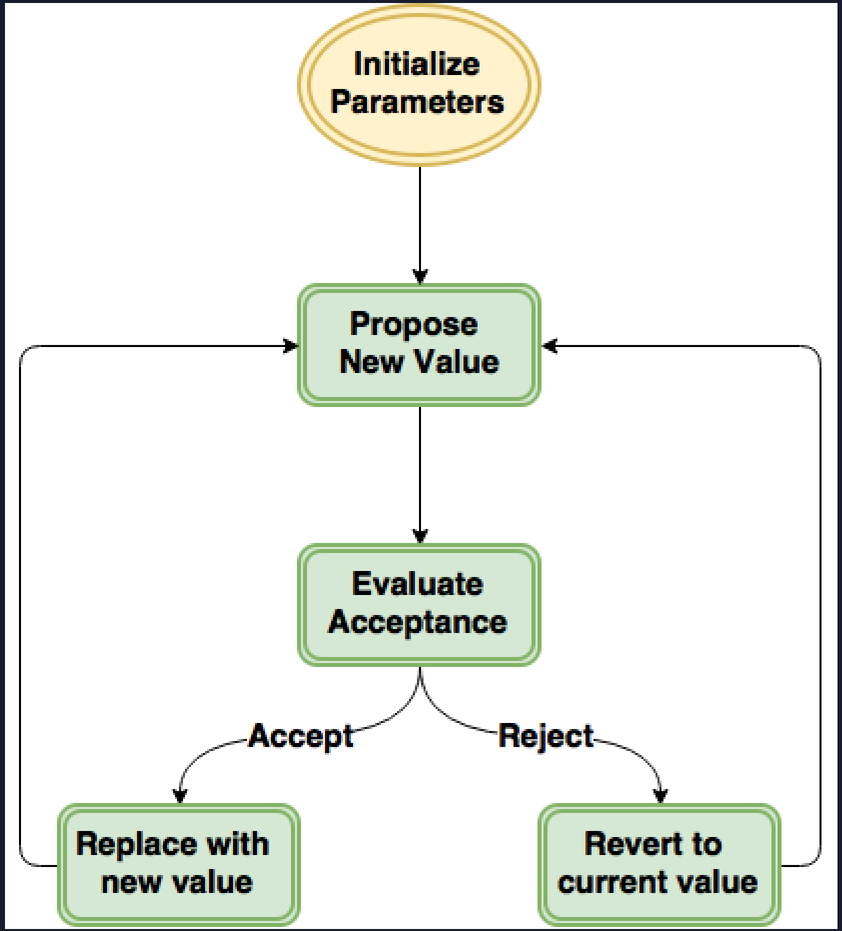
Metropolis sampling

Repeat until convergence:

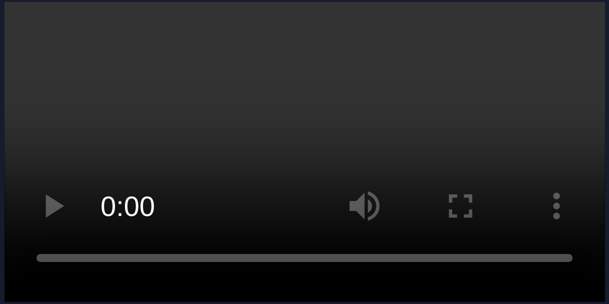
1. Sample θ' from $q(\theta'|\theta^{(t)})$.
2. Generate a Uniform[0,1] random variate u.
3. Calculate acceptance probability:

$$a(\theta', \theta) = \frac{\pi(\theta')}{\pi(\theta)}$$

4. If $a(\theta', \theta) > u$ then $\theta^{(t+1)} = \theta'$, otherwise $\theta^{(t+1)} = \theta^{(t)}$.



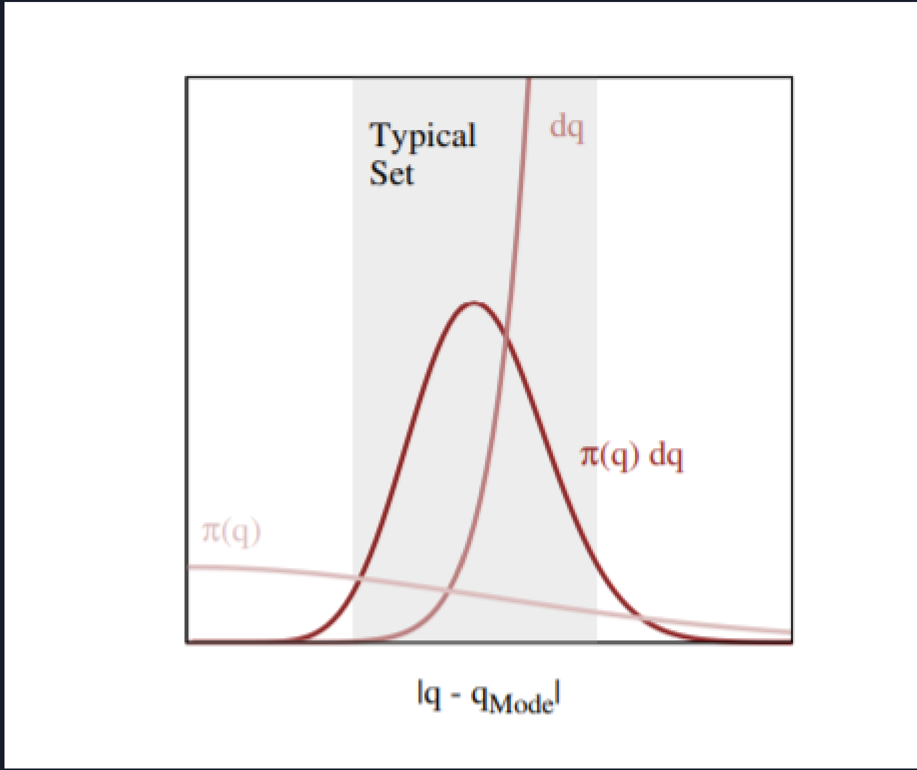
Metropolis sampling



Most of the *probability* of a multivariate distribution lies in the

typical set

rather than near the mode.



Hamiltonian Monte Carlo

Uses a *physical analogy* of a frictionless particle moving on a hyper-surface

Requires an *auxiliary variable* to be specified

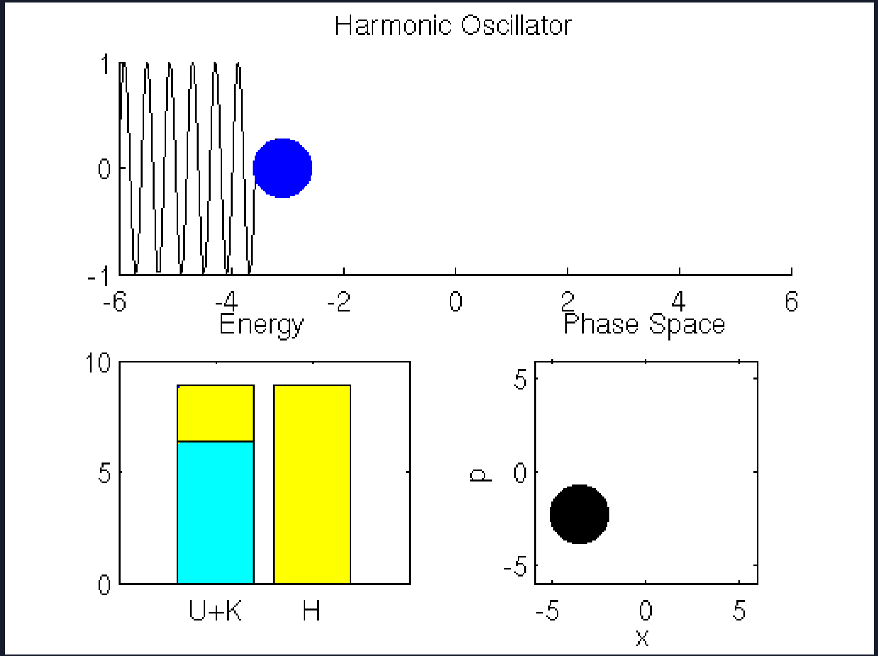
- position (unknown variable value)
- momentum (auxiliary)

$$\mathcal{H}(s,\phi) = E(s) + K(\phi) = E(s) + \frac{1}{2}(\sum_i)\phi_i^2$$

Hamiltonian Dynamics

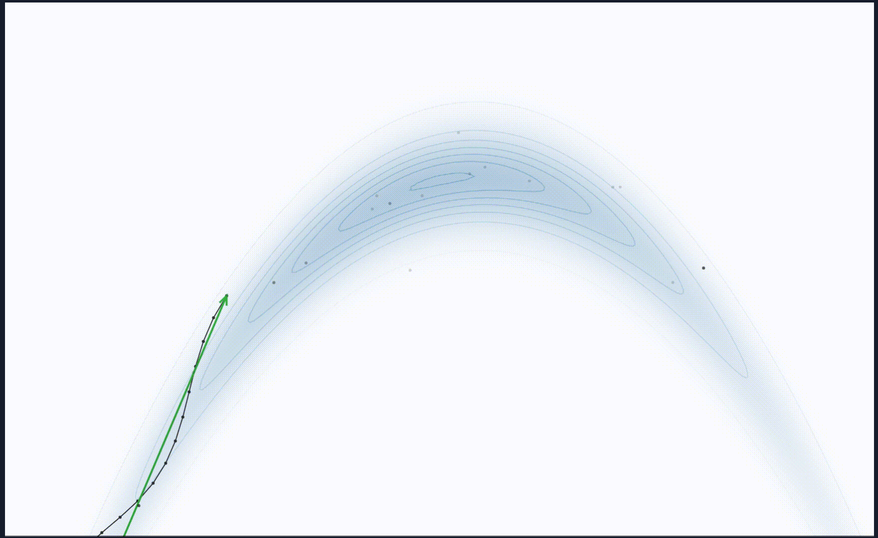
$$\frac{ds_i}{dt} = \frac{\partial \mathcal{H}}{\partial \phi_i} = \phi_i$$

$$\frac{d\phi_i}{dt} = -\frac{\partial \mathcal{H}}{\partial s_i} = -\frac{\partial E}{\partial s_i}$$

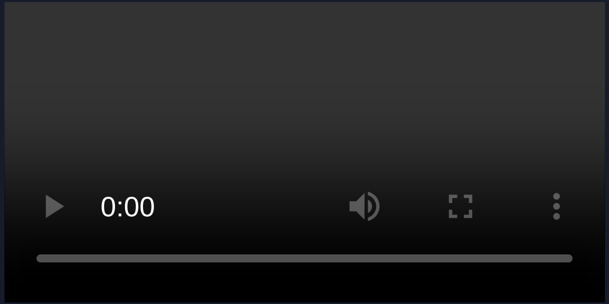


Hamiltonian MC

1. Sample a **new velocity** from univariate Gaussian
2. Perform **n leapfrog steps** to obtain new state θ'
3. Perform **accept/reject** move of θ'

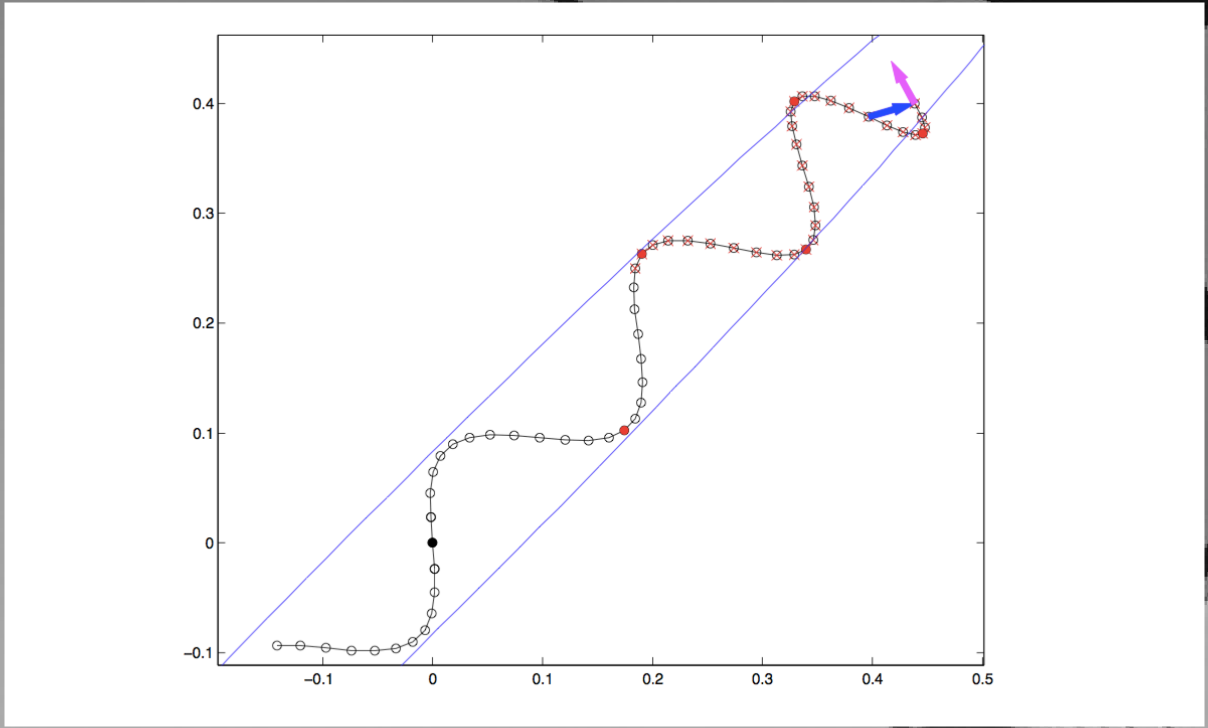


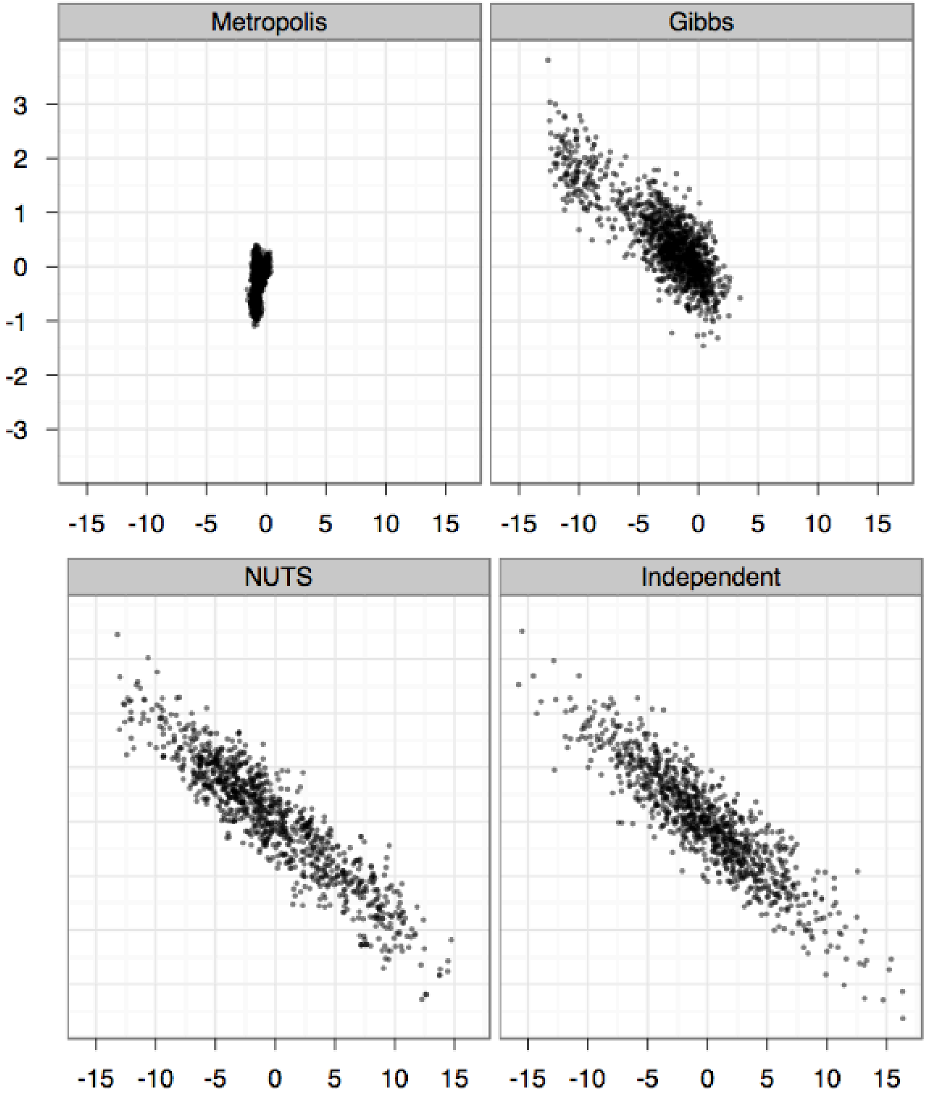
Hamiltonian MC



No U-Turn Sampler (NUTS)

Hoffmann and Gelman (2014)




Metropolis
Gibbs
NUTS
Independent
3
2
1
0
-1
-2
-3
-15
-10
-5
0
5
10
15