Hamiltonian MCMC

Dan Sheldon

April 15, 2024

Review: Metropolis-Hastings

- Given: probability density $P(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^d$
- Goal: generate sample $\mathbf{x} \sim P$

Metropolis-Hastings

- Initialize $\mathbf{x}^{(0)}$ arbitrarily
- Given $\mathbf{x}^{(t)} = \mathbf{x}$, propose

$$\mathbf{x}' \sim Q(\cdot \mid \mathbf{x})$$

- Accept and set $\mathbf{x}^{(t+1)} = \mathbf{x}'$ with probability $\min(a, 1)$

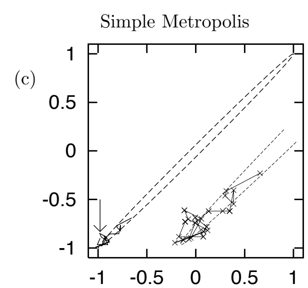
$$a = \frac{P(\mathbf{x}')}{P(\mathbf{x})} \cdot \frac{Q(\mathbf{x} \mid \mathbf{x}')}{Q(\mathbf{x}' \mid \mathbf{x})}$$

- Else reject and set $\mathbf{x}^{(t+1)} = \mathbf{x}^{(t)}$

For large enough T, have $\mathbf{x}^{(T)} \sim P$

Problem: random walk

Slow mixing due to "random walk" behavior



Why? Typical proposal is a random *displacement*

- Spherical Gaussian → Brownian motion-like
- Ignores density surface

Main Idea

- Idea: use density to guide proposals
- Select random *velocity* $\mathbf{p}/m \in \mathbb{R}^d$
  - $\mathbf{p}$ = momentum, m = mass
- Simulate motion on energy surface

$$\{(\mathbf{x}, E(\mathbf{x})) : \mathbf{x} \in \mathbb{R}^d\} \subseteq \mathbb{R}^{d+1}, \quad E(\mathbf{x}) = -\log P(\mathbf{x})$$

with initial velocity $\mathbf{p}/m$ for some amount of time to get proposal $\mathbf{x}'$.

Main Idea

Demo: $P(x)$, $E(x)$, motion in 1D

Hamiltonian Mechanics

- Position $\mathbf{x} \in \mathbb{R}^d$
- Velocity $\mathbf{p}/m \in \mathbb{R}^d$
- Potential energy $E(\mathbf{x})$ (= height)
- Temporal dynamics

$$\frac{d\mathbf{x}}{dt} = \frac{\mathbf{p}}{m}$$
$$\frac{d\mathbf{p}}{dt} = -\frac{\partial E(\mathbf{x})}{\partial \mathbf{x}}$$

Puck of mass m sliding on frictionless surface with velocity $\mathbf{p}/m$, height at $\mathbf{x}$ equal to $E(\mathbf{x})$ (and thus "incline" $\partial E(\mathbf{x})/\partial \mathbf{x}$).

Generalization: Kinetic Energy

Define $K(\mathbf{p}) = \dfrac{\mathbf{p}^T\mathbf{p}}{2m}$

kinetic energy

$$\frac{d\mathbf{x}}{dt} = \frac{\mathbf{p}}{m}, \quad \frac{d\mathbf{p}}{dt} = -\frac{\partial E(\mathbf{x})}{\partial \mathbf{x}} \quad \Longrightarrow \quad \frac{d\mathbf{x}}{dt} = \frac{\partial K(\mathbf{p})}{\partial \mathbf{p}}, \quad \frac{d\mathbf{p}}{dt} = -\frac{\partial E(\mathbf{x})}{\partial \mathbf{x}}$$

Generalization: The Hamiltonian

Define $H(\mathbf{x}, \mathbf{p}) = E(\mathbf{x}) + K(\mathbf{p})$

Hamiltonian or total energy

$$\frac{d\mathbf{x}}{dt} = \frac{\partial K(\mathbf{p})}{\partial \mathbf{p}} \qquad \frac{d\mathbf{p}}{dt} = -\frac{\partial E(\mathbf{x})}{\partial \mathbf{x}} \quad \Longrightarrow \quad \frac{d\mathbf{x}}{dt} = \frac{\partial H(\mathbf{x}, \mathbf{p})}{\partial \mathbf{p}} \qquad \frac{d\mathbf{p}}{dt} = -\frac{\partial H(\mathbf{x}, \mathbf{p})}{\partial \mathbf{x}}$$

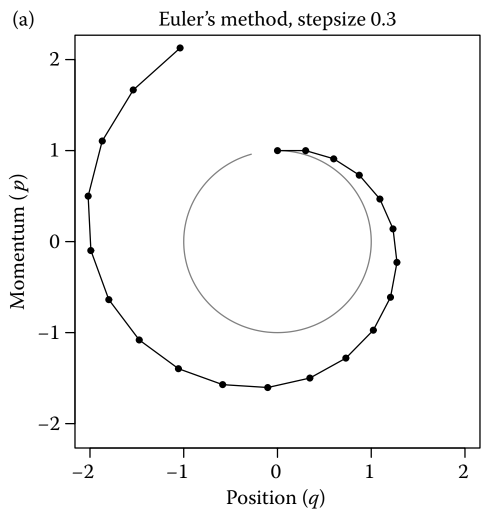
Simluating Hamiltonian Mechanics

Euler's method

$$\mathbf{x}(t+\varepsilon) = \mathbf{x}(t) + \varepsilon \frac{\mathbf{p}(t)}{m}$$
$$\mathbf{p}(t+\varepsilon) = \mathbf{p}(t) - \varepsilon \frac{\partial E(\mathbf{x}(t))}{\partial \mathbf{x}}$$

Problem: numerically unstable

Euler's Method

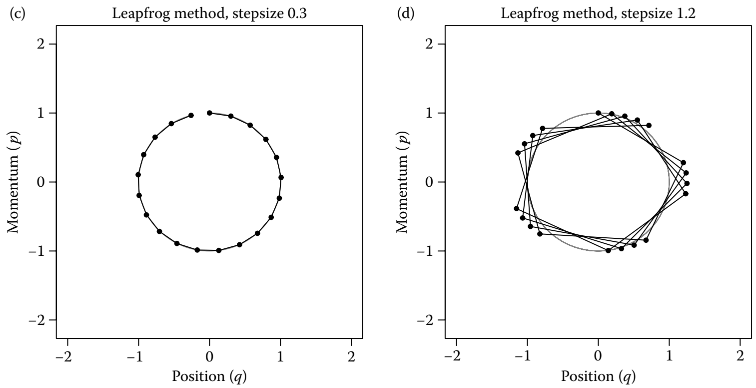


Leapfrog Method

More accurate and stable method

$$\mathbf{p}(t+\varepsilon/2) = \mathbf{p}(t) - (\varepsilon/2)\frac{\partial E(\mathbf{x}(t))}{\partial \mathbf{x}}$$
$$\mathbf{x}(t+\varepsilon) = \mathbf{x}(t) + \varepsilon \frac{\mathbf{p}(t+\varepsilon/2)}{m}$$
$$\mathbf{p}(t+\varepsilon) = \mathbf{p}(t+\varepsilon/2) - (\varepsilon/2)\frac{\partial E(\mathbf{x}(t+\varepsilon))}{\partial \mathbf{x}}$$

Leapfrog Method



Hamiltonian MCMC

Random velocity/momentum instead of random displacement

- Start at $\mathbf{x}$
- Choose random momentum $\mathbf{p} \sim \exp(-\mathbf{p}^T\mathbf{p}/2m)$
- Simluate Hamiltonian mechanics for s time units
  $\longrightarrow$ end at $\mathbf{x}'$
- Propose $\mathbf{x}'$

Problem: how to compute $Q(\mathbf{x}' \mid \mathbf{x})$ for acceptance probability?

Auxilliary Variables

Sample both $\mathbf{x}$ and $\mathbf{p}$ from

$$\begin{aligned} P(\mathbf{x}, \mathbf{p}) &= \exp(-H(\mathbf{x}, \mathbf{p})) \\ &= \exp(-E(\mathbf{x}))\exp(-K(\mathbf{p})), \end{aligned}$$

when done, discard $\mathbf{p}$ values

Note: $\mathbf{x}$ and $\mathbf{p}$ are independent

Hamiltonian MCMC

Gibbs step

- Start at $(\mathbf{x}, \mathbf{p}^-)$
- Choose random momentum $\mathbf{p} \sim \exp(-\mathbf{p}^T\mathbf{p}/2m)$
- End at $(\mathbf{x}, \mathbf{p})$

Metropolis-Hastings step

- Start at $(\mathbf{x}, \mathbf{p})$
- Simluate Hamiltonian mechanics $\rightarrow$ end at $(\mathbf{x}', \mathbf{p}')$
- Propose $(\mathbf{x}', -\mathbf{p}')$

Acceptance Probability?

$$a = \frac{P(\mathbf{x}', -\mathbf{p}')}{P(\mathbf{x}, \mathbf{p})} \cdot \frac{Q(\mathbf{x}, \mathbf{p} \mid \mathbf{x}', -\mathbf{p}')}{Q(\mathbf{x}', -\mathbf{p}' \mid \mathbf{x}, \mathbf{p})}$$

$$= \frac{P(\mathbf{x}', -\mathbf{p}')}{P(\mathbf{x}, \mathbf{p})} \qquad \text{(reversibility, volume preservation)}$$

$$= \exp(E(\mathbf{x}) - E(\mathbf{x}') + K(\mathbf{p}) - K(\mathbf{p}')) \qquad (K(\mathbf{p}') = K(-\mathbf{p}'))$$

$$\approx 1 \qquad \text{(conservation of energy)}$$

Reversibility

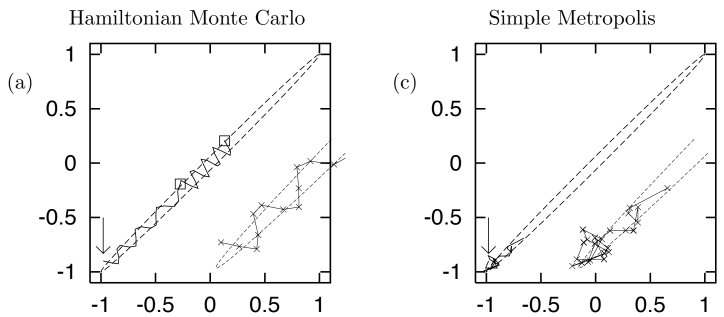
Let $T_{L,\varepsilon} : \mathbb{R}^{2d} \to \mathbb{R}^{2d}$ be simulation mapping with L steps at time increment ε

$$T_{L,\varepsilon}(\mathbf{x}, \mathbf{p}) = (\mathbf{x}', \mathbf{p}') \implies T_{L,\varepsilon}(\mathbf{x}', -\mathbf{p}') = (\mathbf{x}, -\mathbf{p})$$

Demo

Demo with sampling

Example



Example

Setup: 100D Gaussian, standard deviations in different dimensions are $0.01, 0.02, \ldots, 1.00$