Hybrid Monte Carlo

Radford M. Neal. An improved acceptance procedure for the hybrid monte carlo algorithm. *J. Comput. Phys.*, 111(1):194–203, 1994

Swaminathan Sankararaman

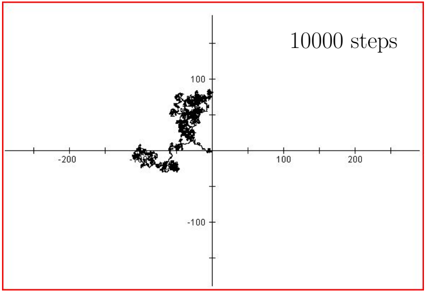
13 Feb 2009

Contents

Current Section

Problem with Naive Metropolis-Hastings

- Random Walks travel expected distance of $O(\sqrt{n})$ after n steps



What to do?

What to do?

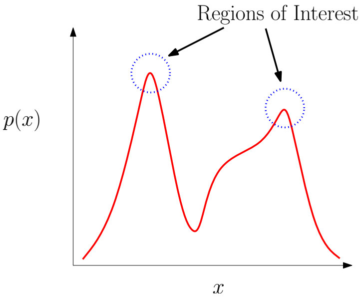
- What do we know about the problem?

What to do?

- What do we know about the problem?
- What do we know about the structure?

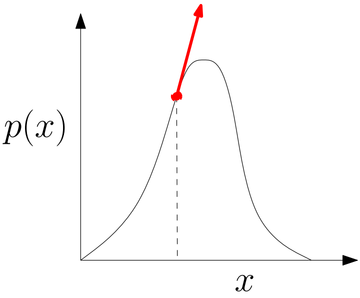
What to do?

- What do we know about the problem?
- What do we know about the structure?



Observation

- When a ball is on an incline, the angle of incline dictates the force with which it rolls down



Characterization for a physical system

- We have a system with particles which have a *position* $\mathbf{x}$ and a *momentum* $\mathbf{u}$

Characterization for a physical system

- We have a system with particles which have a *position* $\mathbf{x}$ and a *momentum* $\mathbf{u}$
- We have energies associated with particles

Characterization for a physical system

- We have a system with particles which have a *position* **x** and a *momentum* **u**
- We have energies associated with particles
  - Potential Energy $E(x)$

$$P(x) = \frac{1}{Z_E}\exp(-E(x))$$

Characterization for a physical system

- We have a system with particles which have a *position* **x** and a *momentum* **u**
- We have energies associated with particles
  - Potential Energy $E(x)$

$$P(x) = \frac{1}{Z_E} \exp(-E(x))$$

  - Kinetic Energy $K(u) = \frac{1}{2} \sum_i u_i^2$

$$P(u) = \frac{1}{Z_k} \exp(-K(u))$$

Characterization for a physical system

- We have a system with particles which have a *position* **x** and a *momentum* **u**
- We have energies associated with particles
  - Potential Energy $E(x)$

$$P(x) = \frac{1}{Z_E}\exp(-E(x))$$

  - Kinetic Energy $K(u) = \frac{1}{2}\sum_i u_i^2$

$$P(u) = \frac{1}{Z_k}\exp(-K(u))$$

- The gradient of $E(x)$ dictates the change in momentum

Characterization for a physical systems

- The total energy of the system is

$$H(x, u) = E(x) + K(u)$$

Characterization for a physical systems

- The total energy of the system is

$$H(x,u) = E(x) + K(u)$$

- The joint distribution over x and u is

$$P(x,u) = \frac{1}{Z_H}\exp(-H(x,u)) = \mathbf{P}(\mathbf{x})\mathbf{P}(\mathbf{u})$$

Characterization for a physical systems

- The total energy of the system is

$$H(x, u) = E(x) + K(u)$$

- The joint distribution over x and u is

$$P(x, u) = \frac{1}{Z_H}\exp(-H(x, u)) = \mathbf{P}(\mathbf{x})\mathbf{P}(\mathbf{u})$$

$P(x)$ and $P(u)$ are independent

Characterization in General

- Target distribution $P(x)$ is everywhere differentiable

Characterization in General

- Target distribution $P(x)$ is everywhere differentiable
- Represent the Potential energy as $\log P(x)$

Characterization in General

- Target distribution $P(x)$ is everywhere differentiable
- Represent the Potential energy as $\log P(x)$
- Augment with Kinetic Energy $K(u)$

Characterization in General

- Target distribution $P(x)$ is everywhere differentiable
- Represent the Potential energy as $\log P(x)$
- Augment with Kinetic Energy $K(u)$
- Sample $P(x, u)$ taking into advantage the gradient.

Characterization in General

- Target distribution $P(x)$ is everywhere differentiable
- Represent the Potential energy as $\log P(x)$
- Augment with Kinetic Energy $K(u)$
- Sample $P(x, u)$ taking into advantage the gradient.
- And Then?

Sampling - Idea

- Use Hamiltonian Dynamics at a given sample (x_i, u_i)

Sampling - Idea

- Use Hamiltonian Dynamics at a given sample (x_i, u_i)

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial u_i} = u$$

Sampling - Idea

- Use Hamiltonian Dynamics at a given sample (x_i, u_i)

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial u_i} = u$$

$$\frac{du_i}{dt} = -\frac{\partial H}{\partial x_i} = -\nabla E(x_i)$$

Sampling - Idea

- Use Hamiltonian Dynamics at a given sample (x_i, u_i)

$$\frac{dx_i}{dt} = \frac{\partial H}{\partial u_i} = u$$

$$\frac{du_i}{dt} = -\frac{\partial H}{\partial x_i} = -\nabla E(x_i)$$

- Modify using above equations to get a new sample

Leapfrog Discretization

$$u(\tau + \frac{\epsilon}{2}) = u(\tau) - \frac{\epsilon}{2} \nabla E(x(\tau))$$

$$x(\tau + \epsilon) = x(\tau) + \epsilon u(\tau + \frac{\epsilon}{2})$$

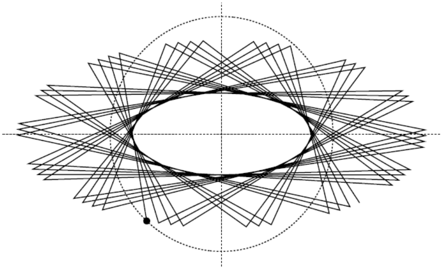
$$u(\tau + \epsilon) = u(\tau + \frac{\epsilon}{2}) - \frac{\epsilon}{2} \nabla E(x(\tau + \epsilon))$$

Time Reversible and Volume Preserving

Current Section

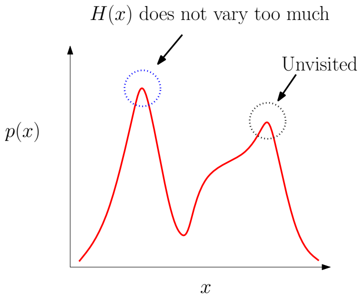
Two problems with the dynamical method

Discretization errors



Two problems with the dynamical method
Energy Well

We might be stuck in an energy well (intuitively, a local maxima)



Stochastic Step

- Replace all values of momentum u with values picked from their distribution

Stochastic Step

- Replace all values of momentum u with values picked from their distribution
- Why can we do this?

Stochastic Step

- Replace all values of momentum u with values picked from their distribution
- Why can we do this?
  - $P(u)$ is independent of $P(x)$

Stochastic Step

- Replace all values of momentum u with values picked from their distribution
- Why can we do this?
  - $P(u)$ is independent of $P(x)$
- What does this get us?

Stochastic Step

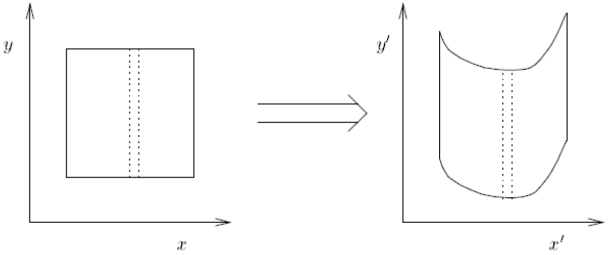
- Replace all values of momentum u with values picked from their distribution
- Why can we do this?
  - $P(u)$ is independent of $P(x)$
- What does this get us?
  - We can move to regions of different energies H
  - Usually enough to ensure ergodicity

Accept/Reject Step

- Similar to MH
- The Acceptance Probability is a function of the difference in energies

$$a(\Delta H) = min(1, exp(-\Delta H))$$

Detailed Balance



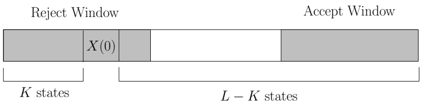
Speed of Computation

- Tunable Parameters - L and ϵ
- Tradeoff between computation time and quality
    - Generally works better for non-local variables (higher dimensions?)

Applications

- Simulation of Physical Systems
- Optimization
- Sampling

Acceptance Procedure using Windows



- Select window to move to
- Select state within window to move to
- Free Energy

$$F(W) = -\log \sum_{X \in W} exp(-H(X))$$

- Acceptance Probability is $a(\Delta F)$

Questions?