Hybrid Monte Carlo

Swaminathan Sankararaman

13 Feb 2009
Contents

1. The Dynamical Method
   - Motivation
   - Solution

2. Hybrid Monte Carlo
   - Motivation
   - How does it work?
   - Analysis
   - Applications
   - Improvement
1. The Dynamical Method
   - Motivation
   - Solution

2. Hybrid Monte Carlo
   - Motivation
   - How does it work?
   - Analysis
   - Applications
   - Improvement
Problem with Naive Metropolis-Hastings

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

![Diagram showing random walks after 10,000 steps]
What to do?

What do we know about the problem?

What do we know about the structure?

 Regions of Interest

 $x$

 $p(x)$

 Regions of Interest
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* $x$ and a *momentum* $u$. 
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
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

\[ H(x, u) = E(x) + K(u) \]
The total energy of the system is

\[ H(x, u) = E(x) + K(u) \]

The joint distribution over \( x \) and \( u \) is

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

\( P(x) \) 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?
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 \((x_i, u_i)\)

\[
\frac{dx_i}{dt} = \frac{\partial H}{\partial u_i} = u
\]

\[
\frac{du_i}{dt} = -\frac{\partial H}{\partial x_i} = -\nabla E(x_i)
\]
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
1. The Dynamical Method
   - Motivation
   - Solution

2. Hybrid Monte Carlo
   - Motivation
   - How does it work?
   - Analysis
   - Applications
   - Improvement
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)

\[ H(x) \text{ does not vary too much} \]

\[ p(x) \]

\[ x \]
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 $P(x)$

$P(u)$ 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

The Dynamical Method
Hybrid Monte Carlo

Motivation
How does it work?
Analysis
Applications
Improvement

Swaminathan Sankararaman
Hybrid Monte Carlo
Tunable Parameters - $L$ and $\epsilon$

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?