Lecture 22: # Hybrid Monte Carlo.

Note Title 5/23/2006

Want to avoid the "random walk" behavior that is common to Metropolis type algorithm.

Motivated by Physics

Newtonian Physics:

$\underline{x}(t)$ d-dim position vector of a particle at time t. $\underline{x} = (x_1, \ldots x_d)$

d-dim mass vector. $\underline{m} = (m_1, \ldots m_d)$

$\underline{v}(t) \triangleq \dot{\underline{x}}(t)$ is the velocity vector. $\left(\dot{x}(t) = \frac{dx}{dt}\right)$

$\dot{\underline{v}}(t)$ is the acceleration

$\underline{F} = \underline{m}\,\dot{\underline{v}}(t)$ Newton's Laws of Motion.

$\underline{m}\,\underline{v} = (m_1 \dot{v}_1, \ldots, m_d \dot{v}_d)$

Momentum $\underline{p} = \underline{m}\,\underline{v}$.

Kinetic Energy $K(\underline{p}) = \frac{1}{2}\sum_{i=1}^{d} m_i v_i^2 = \frac{1}{2}\sum_{i=1}^{d} \frac{p_i^2}{m_i}$

Total energy: $\mathcal{H}(\underline{x}, \underline{p}) = \mathcal{U}(\underline{x}) + K(\underline{p})$

potential energy

$\underline{F} = -\underline{\nabla}\,\mathcal{U}(\underline{x})$

page 2.

Hamilton's equations of motion.

$$\underline{\dot{x}}(t) = \frac{\partial H}{\partial \underline{p}}(\underline{x}:\underline{p})$$

$$\underline{\dot{p}}(t) = -\frac{\partial H}{\partial \underline{x}}(\underline{x}:\underline{p})$$

Note: $\frac{dH}{dt} = \frac{\partial H}{\partial \underline{p}} \cdot \frac{d\underline{p}}{dt} + \frac{\partial H}{\partial \underline{x}} \cdot \frac{d\underline{x}}{dt} = 0$

Energy is conserved.

Discretize the equations:

Finite Differences:

Leapfrog Method:

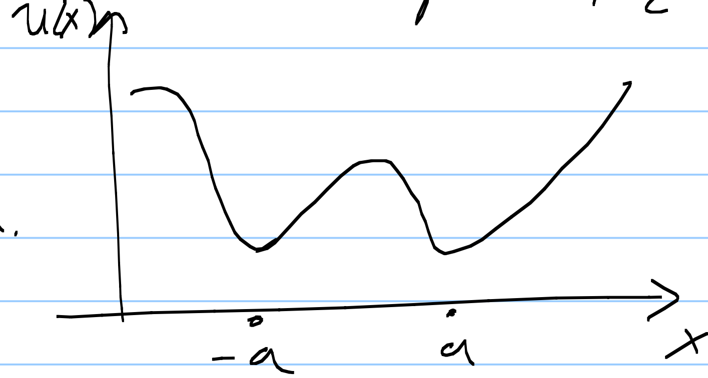
$$x(t+\Delta t) = x(t) + \Delta t \cdot \frac{p}{m}(t+\tfrac{1}{2}\Delta t)$$

$$p(t+\tfrac{1}{2}\Delta t) = p(t-\tfrac{1}{2}\Delta t) + \Delta t \cdot \frac{\partial H}{\partial x}\Big|_t$$

(Momentum at time t is $\frac{1}{2}\{p(t+\frac{1}{2}\Delta t) + p(t-\frac{1}{2}\Delta t)\}$)

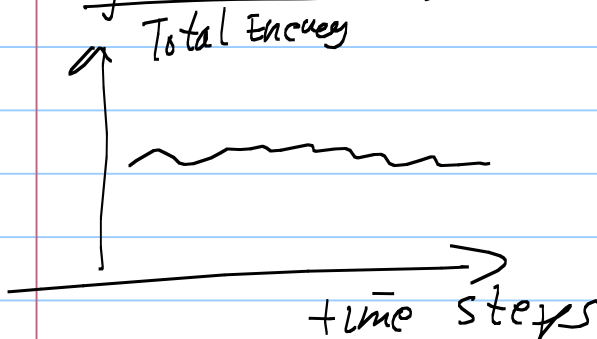
Example: $U(x) = x^2 + a^2 - \log(\cosh(ax))$
$K(p) = p^2/2$

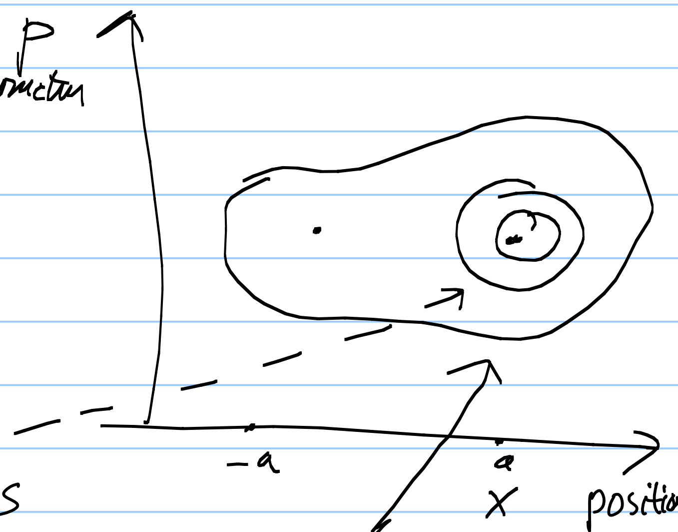
"Two wells" at $x = \pm a$.



Two minima (equal height) at $x = \pm a$.

After Discretization: energy is not preserved.





If initial speed is small, then particle stays in one well.

If initial speed is big, then particle travels between two wells.

Page 4.

Volume Preservation: In leap-frog, as in exact Hamiltonian Dynamics, the volume is preserved from one step to the next.

Why do we care?

<u>Volume</u> $$|V(t)| \stackrel{\text{def}}{=} \iint_{V(t)} dx\,dp.$$

where $V(t) = \{(x(t), p(t)) : (x(0), p(0)) \in V(0)\}$,

Volume preservation: $|V(t)| = |V(0)|$.

Intuiton for Hybrid Monte Carlo.

The momentum allows the particle to escape from local minima.

And the gradient helps guide the particle in the right directon

Hybrid Monte Carlo uses the basic idea of Hamiltonian Systems together with the Metropolis acceptance rule.

Page 5.

Two Basis Observations:

(a) If we can sample from a distribution $\pi(\underline{x}, \underline{p}) \propto e^{-\mathcal{H}(\underline{x}, \underline{p})}$ then we can sample from the marginals $\underline{x} \sim \pi(\underline{x}) \propto e^{-U(\underline{x})}$, $\underline{p} \sim \phi(\underline{p}) \propto e^{-p^2/2m}$.

(b) The Hamiltonian trajectory is time reversible ~ relates to volume preserving.

If we run leap-frog from $(\underline{x}, \underline{p})$ to $(\underline{x}', \underline{p}')$ in t steps, then we can start from $(\underline{x}', -\underline{p}')$ and run t steps to get $(\underline{x}, -\underline{p})$.

This is needed to ensure reversibility (detailed balance) of the Hybrid Monte Carlo.

Page 6.

Hybrid Monte Carlo

Note: the kinetic energy $K(\underline{p})$ is quadratic in $\underline{p}$. This means that $e^{-K(\underline{p})}$ is a Gaussian distribution.

At time t and position $\underline{x}$.

- Generate a new momentum vector $\underline{p}$ from the Gaussian $\phi(\underline{p}) \propto e^{-K(\underline{p})}$.
- Run the leap-frog algorithm starting from $(\underline{x}, \underline{p})$ for L steps, to obtain a new state $(\underline{x}', \underline{p}')$
- Accept the proposed state $(\underline{x}', \underline{p}')$ with probability

$$\min\{1, \exp\{-\mathcal{H}(x', p') + \mathcal{H}(\underline{x}, \underline{p})\}\}$$

otherwise do nothing. Note: time reversibility ensures that the proposal distribution is symmetric. So cancels in the acceptance propability.

Why does this work? Heuristic Argument.

$g^L(\cdot,\cdot)$ is L steps of the leapfrog algorithm.

(a) Let $(\underline{x}', p') = g^L(\underline{x}, \underline{p})$

$(\underline{x}, -\underline{p}) = g^L(\underline{x}', -\underline{p}')$

(b) $\pi(\underline{x}, \underline{p}) = \pi(\underline{x}, -\underline{p})$ for any $\underline{x}, \underline{p}$

(c) By volume preservation $d\underline{x}\, d\underline{p} = d\underline{x}'\, d\underline{p}'$

Suggests the proposal is symmetric (as required by Metropolis). More rigorous proof in Liu.