Monte Carlo Simulation

Vignesh M Pai (20211132)

1 Statistical Physics

1.1 Hamiltonian Dynamics

A dynamical system for our purposes is a space $X = Q \times P$ along with a Hamiltonian $H : X \to \mathbb{R}$ and an initial state (q_0, p_0). The dynamics are given by the first order equations

$$\dot{q}_i = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = -\frac{\partial H}{\partial q_i}$$

We call the parametrized curve $(q(t), p(t))$ that solves these equations with the initial state given, the trajectory of the dynamical system. $(q(t), p(t))$ is also called the state of the system at time t.

We know that Hamiltonian dynamics preserve the Hamiltonian (energy), therefore given an initial state with a fixed energy, we define the phase space of this system to be the set of all points in X with this given energy (the surface $H(q, p) = H(q_0, p_0)$).

1.2 Microcanonical Ensemble

Consider an ensemble (large collection) of dynamical systems differing only perhaps in the initial state, we define the fluid of the ensemble at a fixed time to be the set of all states of systems in the ensemble. Liouville's theorem states that the measure of this fluid is conserved, it directly follows from this that the 'density' (or probability density) of the fluid (assuming a continuous fluid) is locally conserved.

A microcanonical ensemble is an ensemble of dynamical systems whose initial fluid state has uniform density over a surface of constant energy (and zero density everywhere else). Liouville's theorem clearly implies that the fluid of such a system is time independent and thus averaging over the ensemble is the same averaging over the constant energy surface. Further, properties of this ensemble are only dependent on the energy of the system.

1.3 Ergodic Theory

In order for the microcanonical ensemble and its results to be useful, we need to show that the time average of a dynamical system with an arbitrary initial state approaches the microcanonical ensemble average, this comes under ergodic theory. We define an ergodic component of the phase space as an equivalence class on the phase space, two points lie on the same ergodic component if there exists a trajectory connecting them.

We define a dynamical system to be ergodic, if there exists an ergodic component that is dense in phase space. Equivalently, if the ergodic components with non zero volume have volume equal to that of phase space. For a dynamical system with initial state (q_0, p_0), the time average of a quantity O is defined as

$$\bar{O}(q_0, p_0) := \lim_{T \to \infty} \frac{1}{T} \int_0^T O(q(t), p(t)) dt$$

Clearly $\bar{O}$ is constant over a trajectory, hence it is constant over an ergodic component, and thus over phase space except a region of zero volume. This shows the required result.

1.4 Entropy

1.4.1 Parametrization of Entropy

Let $\mathcal{D}$ be the set of all dynamical systems. We define the entropy $S : \mathcal{D} \to \mathbb{R}$ as the logarithm of the volume of the phase space of the dynamical system. Let U be an open set in $\mathbb{R}^k$, we define the parametrization of $\mathcal{C} \subset \mathcal{D}$ as a bijective map $\phi : U \to \mathcal{C}$. If $\mathcal{C}$ has a parametrization $\phi : U \to \mathcal{C}$, the entropy S can be viewed as a map $S : U \to \mathbb{R}$ redefined as $S := S \circ \phi$.

1.4.2 Additive Functions

Let A, B be dynamical systems with Hamiltonians $H_A : X_A \to \mathbb{R}, H_B : X_B \to \mathbb{R}$. The union of the systems $C = A \cup B$ is defined as the system with Hamiltonian $H_C : X_A \times X_B \to \mathbb{R}$ defined as

$$H_C(x_A, x_B) = H_A(x_A) + H_B(x_B)$$

A, B are said to be subsystems of C. Let $\mathcal{D}' \subset \mathcal{D}$ be closed under unions. A function $f : \mathcal{D}' \to \mathbb{R}^k$ is called additive if

$$f(A) + f(B) = f(A \cup B), \quad \forall A, B \in \mathcal{D}'$$

Define an additive parametrization as a parametrization of entropy ϕ such that ϕ^{-1} is additive.

1.4.3 Interacting Systems

Let C_I be a dynamical system with state space X and state given by $x(t)$. A subsystem decomposition of C_I is pair of functions $\Psi_1 : X \to \mathcal{C}_1$ and $\Psi_2 : X \to \mathcal{C}_2$ where $\mathcal{C}_1, \mathcal{C}_2$ are subsets of $\mathcal{D}$ with parametrizations ϕ_1, ϕ_2 (of the same dimension) such that the function

$$G := \phi_1^{-1} \circ \Phi_1 + \phi_2^{-1} \circ \Phi_2$$

is approximately $\langle G \rangle$ with high probability (measured uniformly over X).

Our interest is in finding $\phi_i^{-1} \circ \Phi_i$.

Suppose that $E = \phi^{-1}(\Phi_1(x))$, then the number of states x with this property is

$$\exp(S(E)) \exp(S(E_0 - E))$$

which is maximized for

$$\boldsymbol{\nabla} S(E) =$$

1.4.4 Canonical Ensemble

2 Metropolis Hastings Algorithm

Suppose ρ be a probability distribution over a space X and suppose $r : X \times X \to \mathbb{R}$ is the relative probability function

$$r(x, y) = \frac{\rho(x)}{\rho(y)}$$

In practice, it is often much easier to compute r than ρ. The Metropolis Hastings algorithm is a Markov chain algorithm that generates samples from X with distribution ρ, given r.

A Markov chain is a sequence of random numbers that are generated iteratively by sampling from a transition probability distribution T. Concretely, if x_k is the kth element in the sequence, x_{k+1} is chosen with probability $T(x_{k+1}, x_k)$. Our goal is to find T such that the sequence of random numbers eventually follow ρ. A sufficient condition for this is known as the principle of detailed balance

$$T(y, x)\rho(x) = T(x, y)\rho(y), \quad \forall x, y \in X$$

intuitively, this means that there is equal probability to transition between any two points in the samples of the distribution ρ.

The algorithm splits the transition into two step: proposal and acceptance. In the proposal step, a point x_p is proposed to be the $(k+1)$th element by sampling from the distribution $g(x_p, x_k)$, this is usually a function that has high probability for x_p 'close' to x_k. In the acceptance step, we accept the proposed point x_p with probability $A(x_p, x)$. Substituting back to the detaile balance equation, we get

$$A(y,x)g(y,x)\rho(x) = A(x,y)g(x,y)\rho(y)$$
$$\implies \frac{A(y,x)}{A(x,y)} = r(y,x)\frac{g(x,y)}{g(y,x)}$$

We have reduced the problem of finding T to finding A. In this algorithm, we choose A to be defined as

$$A(y,x) = \min\left(1, r(y,x)\frac{g(x,y)}{g(y,x)}\right)$$

3   Ising Model

The Ising model is a d dimensional lattice of length $L-1$ in each dimension with a spin at each coordinate that may take values ± 1. The state of the system is completely determined by the values of the spins, for a microstate S, $S_{[i]}$ denotes the value of the spin at coordinate $[i]$. The Hamiltonian of the microstate S can be written as

$$H(S) = -\sum_{[i],[j]} S_{[i]} S_{[j]}$$

where the summation is over nearest neighbors $[i], [j]$.

The partition function and thus the pdf for this model is difficult to compute, however r given by

$$r(S_1, S_2) = \exp\left(-\frac{H(S_1) - H(S_2)}{k_B T}\right)$$

is easy to compute especially when S_1 is 'close' to S_2. Therefore, we simulate with Metropolis Hastings with g given by:

$$g(S_p, S) = \begin{cases} \frac{1}{L^d} & S_p \text{ differs from } S \text{ at exactly 1 point} \\ 0 & \text{otherwise} \end{cases}$$

Clearly this g is symmetric $(g(x,y) = g(y,x))$, therefore A simplifies to

$$A(y,x) = \min(1, r(y,x))$$