Monte Carlo Simulation

Vignesh M Pai (20211132)

1 Statistical Physics

1.1 Mathematical Formalism

A dynamical system for our purposes is a phase 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 that solves these equations with the initial state given, the trajectory of the dynamical system. For present purposes, we do not compare dynamical systems of different Hamiltonians, assume the following results for a fixed Hamiltonian.

Consider an ensemble (large collection) of dynamical system 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 system 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.

In order for the microcanonical ensemble and its results to be useful, we need to show that an arbitrary ensemble approaches the microcanonical ensemble. Birkhoff's theorem states (roughly) that if we have a function $f : X \to \mathbb{R}$, then for a dynamical system on X, the time average of this function is well defined. Further, this time average is constant over all dynamical systems whose initial state lies on the same trajectory.

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

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