Teach your cat Monte Carlo methods- Part 1

Shraddha Agrawal

Abstract. These are my notes on Monte-Carlo methods. These methods are useful in integration problems, with several applications in many-body physics calculations. These notes are based on the NPTEL lecture series on "Advanced Thermodynamics and Molecular Dynamics" by Prof. Prateek Jha and from the book "Computational Physics" by Prof. JM Thijssen. The Jupyter notebooks will complement these lecture notes in the repository.

If you find any typos or errors, please email me at sa30@illinois.edu.

Table of contents

1 Introductory remarks about Monte Carlo methods

Every numerical technique where random numbers play a key role can be called a Monte Carlo technique. It is used in physics to study classical and quantum many-body particle systems. Outside of physics, it is also used as a way of calculating high-dimensional integrals. Based on applications and operating principles, Monte Carlo can usually be classified into three different categories.

The first category is Direct Monte Carlo where random numbers are used to model physical phenomena, for example, traffic jams.

The second category is Monte Carlo integration which is a method to calculate integrals using random numbers. We will give a heuristic argument as to why the Monte Carlo technique is better for higher dimensional integrals in section ..

The third category is the Metropolis Monte Carlo, in which a sequence of operations on a system is generated in a Markov Chain. The details and the code will follow in the next set of notes, where we will discuss the application of Metropolis Monte Carlo for a 2D Ising system and a monoatomic ideal gas.

2 Monte Carlo versus molecular dynamics

The phase space-space of any system of N particles is a $6N$ dimensional space, where any one point in the phase-space represents one of the states of the system. The goal, in general, is to predict various quantities of interest, for example, the equilibrium magnetization of a system with the next nearest ferromagnetic interaction and a transverse field, or the pressure of a mono-atomic ideal gas. Sampling all possible states for 10^{23} particles might be hard on a computer. Essentially, two approaches sample a restricted set of states while still predicting various phenomena. These two approaches are molecular dynamics simulations and Monte Carlo simulations.

In molecular dynamics, we follow the trajectory $\Gamma(t)$ in the phase space. While molecular dynamics is not the subject of discussion in these notes, it is worth noting that molecular dynamics is great for calculating dynamic properties of the system like transport of heat or charge, or more importantly *systems far from equilibrium*.

The Monte Carlo method, on the other hand, is a random sampling method primarily for equilibrium systems. It is a purely mathematical technique used for integration and has applications far beyond physics. The Monte Carlo method can be extended to non-equilibrium systems, more on this later.

Since the Monte Carlo method is a sampling method for systems at equilibrium, the next section introduces the idea of equilibrium for an N-particle system. We will need these ideas later on to understand the Metropolis algorithm.

2.1 Detour 1-Liouville's theorem

Liouville's theorem states that the phase space density $\rho(\Gamma, t)$ for an N particle system can be thought of as an incompressible fluid. Here $\Gamma = \prod_{i=1}^{N} d^3p_i d^3q_i$ is an infinitesimal volume around the point $(\mathbf{p}, \mathbf{q})$. The proof for this theorem can be found in multiple books (Mehran Kardar's notes on statistical physics are an excellent reference).

In this treatment, we denote the probability density by $\rho(\mathbf{p}, \mathbf{q})$ and it gives us the probability $\Gamma\rho(\mathbf{p}, \mathbf{q})$ that the system will be found in the infinitesimal volume Γ.

Irrespective of equilibrium or not, because of Liouville's theorem, $\frac{d\rho}{dt} = 0$. The total derivative with respect to time can be broken down into the following

terms-

$$\frac{\partial}{\partial t} + \dot{p}.\nabla_q + \dot{q}.\nabla p \tag{1}$$

Note the distinction between $\frac{\partial \rho}{\partial t}$ and $\frac{d\rho}{dt}$. The former partial derivative refers to the changes in ρ at a particular location in phase space, while the latter total derivative follows the evolution of a volume of fluid as it moves in phase space.

2.2 Detour 2-The condition for equilibrium

The demand that phase space behaves like an incompressible fluid gives us the condition that $\frac{d\rho}{dt} = 0$. The demand for *equilibrium*, on the other hand, will impose an additional condition that $\frac{\partial \rho}{\partial t} = 0$ as well. To see this, we note that the average for a variable $O(\mathbf{p}, \mathbf{q})$ is given by the integral of the variable times the probability density-

$$< O >= \int d\Gamma \rho(\mathbf{p}, \mathbf{q}) O(\mathbf{p}, \mathbf{q}) \tag{2}$$

In equilibrium, all averages should be independent of time, and so the partial derivative of ρ should be independent of time as well.

3 Hello world version of Monte-Carlo

The "Hello World" version of Monte Carlo is finding out the value of π. Here the idea is that we sample randomly in the region formed by the coordinates- (-1,-1), (-1,1), (1,-1), and (1,1). We then calculate the number of points that are inside the circle with center (0,0) and radius=1 and divide it by the number of points inside the square formed by (-1,-1), (-1,1), (1,-1), and (1,1). The ratio of these two points is π. Effectively, we integrated- $\int_{-1}^{+1} \sqrt{1-x^2}$. The code for this calculation is uploaded to the repository.

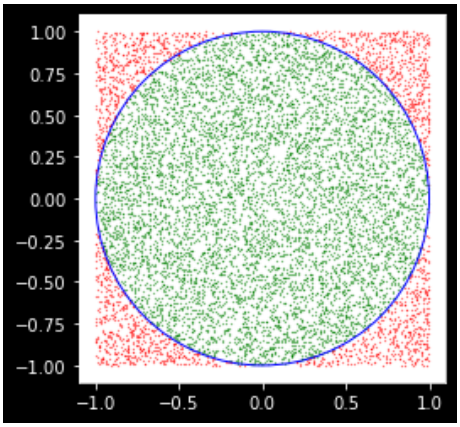

Image 1: Calculating the value of π using random sampling. In the code used to generate this image, we randomly sampled 10000 points. The difference between the real and calculated value of π is 7×10^{-6}

4 The real advantage of Monte Carlo?

As we hinted in the last section, the utility of Monte-Carlo arises because of random sampling. This random sampling proves to be computationally better for integrals in higher dimensions. For physics problems, Monte Carlo means sampling the phase space by looking at many points in the phase space. To see why Monte Carlo methods are better at performing integrals in higher dimensions, consider the following integral-

$$I = \int_a^b f(x) dx \tag{3}$$

Where $f(x)$ is a smooth function over the interval $[a, b]$. With the usual quadrature integration methods, the integration process entails calculating the value of the function over equidistant values of x_i and then evaluating the sum-

$$I = \frac{(b-a)}{N} \sum_{i=1}^{N} w_i f(x_i) \tag{4}$$

Where w_i's are weights that determine the accuracy of the integration method. In Monte Carlo integration, on the other hand, the x_i's are sampled randomly and w_i's are all set to 1. The key difference is how the accuracy of these integrals depends on the number of sampling points.

The accuracy of the quadrature methods will depend will be proportional to N^{-k} or h^k, where N is the number of points sampled, h is the spacing between

the points being sampled, and k is a positive integer. For a d dimensional integral, the number of points sampled N will be equal to $(\frac{L}{h})^d$, where L is the typical length associated with the integration volume. The error for this d-dimensional integral will, therefore scale as $N^{\frac{-d}{k}}$. For Monte Carlo methods, on the other hand, the error will depend on the number of points as $N^{\frac{-1}{2}}$ (following the central limit theorem). Thus for $d \geq 2k$, Monte Carlo will outperform the usual numerical integration procedures.

We can sample any points in the phase space, with some rules of course. Those rules need not reflect any particular set of laws, such as Newton's laws of motion. They can be randomly sampled, or sampled with some "importance" rule.

In the next set of notes, we will discuss importance sampling, and the metropolis algorithm in detail.