

1 Monte Carlo and the Metropolis Algorithm

The Monte Carlo method is a broad class of computational algorithms where the use of random sampling yields numerical results. Any problem that has a probabilistic interpretation can be solved with the Monte Carlo method. And, due to the law of large numbers, with a sufficiently large sample size, expectation values (or, analogously, any integral) can be approximated by a simple mean. For example, we can approximate the following integral:

$$\int_a^b f(x) dx = \frac{b-a}{N} \sum_{i=1}^N f(x_i) \quad (1)$$

for N samples. This integral can be generalized to any finite dimensional integral over some volume Ω as

$$\int_{\Omega} f(x) dx = \frac{V(\Omega)}{N} \sum_{i=1}^N f(x_i) \quad (2)$$

where $V(\Omega)$ is the volume of Ω . This is valid for any dimension and has little effect on computational time. In this paper, high dimensional integrals are computed numerically and so this method holds an advantage.

1.1 Markov Chain

A Markov chain is a sequence of states (x_0, x_1, \dots, x_n) such that any state x_{i+1} depends only on its immediately preceding state x_i , a property called memorylessness. The probability of the transition from state x_i to state x_{i+1} is the transition amplitude and defining such a function allows for the modeling of a Markov chain. Every model in this paper has the property of memorylessness and can be modeled as a Markov chain.

Moreover, every model can be modeled by Markov chain Monte Carlo methods. The equilibrium of a many-state system can be difficult, if not impossible, to solve. Rather, starting with an initial state x_0 , the system will fall into equilibrium, or reach the target distribution, at state x_n for some n large enough. The initial state x_0 is arbitrary although a guess of the equilibrium distribution is helpful so as to converge quicker. Beforehand, the equilibrium distribution is unknown but it can be approximated by evolving the Markov chain until the resulting distribution satisfies within an error the conditions of the equilibrium state. As stated before, there is a transition amplitude between states x_i and x_{i+1} . So, for some change on x_i , a decision must be made whether it will change to x_{i+1} or not change. The method used in this paper is called the Metropolis algorithm.

1.2 Metropolis Algorithm

For a Markov chain, there is a probability density $f(x_j|x_i)$ of choosing state x_j as the next candidate in the Markov chain given state x_i . The change in x_j from x_i is usually small. This is to allow for a large number of accepted changes but it isn't too small so as to not change the state negligibly. There is also the conditional probability $P(x_j|x_i)$ of changing to state x_j given state x_i . If we assume the process is reversible, then it satisfies the condition of detailed balance

$$q(x_j|x_i)P(x_i) = q(x_i|x_j)P(x_j) \quad (3)$$

where $P(x)$ is the desired probability distribution and $q(x_j|x_i)$ is the transition amplitude from x_i to x_j and noting that $q(x_j|x_i) = f(x_j|x_i)P(x_j|x_i)$, then this can be rewritten as

$$\frac{P(x_j|x_i)}{P(x_i|x_j)} = \frac{f(x_i|x_j)q(x_j)}{f(x_j|x_i)q(x_i)}. \quad (4)$$

We wish to find the probability $P(x_j|x_i)$ that satisfies equation 4. The Metropolis choice is

$$P(x_j|x_i) = \min \left(1, \frac{f(x_i|x_j)P(x_j)}{f(x_j|x_i)P(x_i)} \right). \quad (5)$$

Thus, the probability of changing from state x_i to state x_j is shown in equation 5. The Metropolis steps are as follows:

1. Initialize the system with some initial state x_0 .
2. Randomly pick a state x_j with probability given by $f(x_j|x_i)$ where x_i is the current state.
3. Accept this new state with probability $P(x_j|x_i)$ as given in equation 5.
4. If accepted, change the state from x_i to x_j .
5. Otherwise, keep the state x_i .
6. Repeat steps 2-5.

For the models in this paper, a uniform distribution is used for choosing a new state x_j , thus $f(x_j|x_i) = f(x_i|x_j)$. Equation 5 can be rewritten as

$$q(x_j|x_i) = \min \left(1, \frac{P(x_j)}{P(x_i)} \right). \quad (6)$$

From the above equation, we can see that if the probability of state x_j is equal to or greater than the probability of x_i (i.e. the new state is more likely in the equilibrium distribution than the old state), then the change is automatically accepted. Thus acceptance of a new state occurs with certainty when $P(x_{new}) \geq P(x_{old})$. This can once more be rewritten in a piecewise fashion as

$$P(x_{i+1}|x_i) = \begin{cases} 1 & \text{if } \frac{P(x_{i+1})}{P(x_i)} \geq 1 \\ \frac{P(x_{i+1})}{P(x_i)} & \text{otherwise.} \end{cases} \quad (7)$$

The functions for $P(x_i)$ will be defined later in the paper for the specific models.

The Metropolis algorithm allows for the evolution of a statistical system without necessarily the knowledge of the desired probability distribution. But Monte Carlo can also be used in the measurement of observables of the system; after all, not much use is gained by modeling a system if no measurements are made on it.

1.3 Monte Carlo Measurements

As shown in equation 2, Monte Carlo has its uses in the numerical integration of multi-dimensional integrals. We will see that the latter two models in this paper are derived from the Feynman path integral which is, essentially, an infinite dimensional path integral which accounts for every possible path from one state to another (since in quantum mechanics, every path will have a nonzero probability due to effects such as tunneling).

When discretized, any integral is approximated by a sum. We will see that for some expectation average $\langle f(x) \rangle$ that is weighted by some probability density function $g(x)$, it can be written as

$$\langle f(x) \rangle = \int f(x)g(x) dx. \quad (8)$$

But if some x has a probability of occurring of $P(x) \propto g(x)$, then the unweighted, simple average

$$\langle f(x) \rangle \approx \overline{f(x)} = \frac{1}{N} \sum_{i=1}^N f(x_i) \quad (9)$$

can be used as an approximation. The term $\overline{f(x)}$ is the Monte Carlo estimator for $\langle f(x) \rangle$. Since N is finite, the estimator will never be exact. In the limit, $\overline{f(x)} \rightarrow \langle f(x) \rangle$ as $N \rightarrow \infty$. This result allows for the approximation of any expectation value where the error is dependent on N . Any probability can be written in terms of an expectation value and thus observables can be measured with this transformation and equation 9.