Monte Carlo

Monte Carlo (MC) comes in two flavors:

① Equilibrium or "static" MC, which is evaluating expectation values of macroscopic observables in the canonical ensemble:

$$E[\psi] = \int \psi(z) \frac{e^{-\mathcal{H}(z)/k_B T}}{\mathcal{Z}} dz$$

② Non-equilibrium or "dynamic" MC, which is simulating a dynamics (at a coarse-grained level) described by a discrete or continuous Markov chain

Monte Carlo numerical integration is an alternative to deterministic quadrature in which one simply evaluates the integrand at a number of sample points:

$$\int_{\Omega \subseteq \mathbb{R}^n} f(x)\,dx \approx |\Omega|\, E[Y] = |\Omega| \left[\frac{1}{N}\sum_{k=1}^{N} f(X_k)\right] = \hat{Y}$$

where $X_k \in \Omega$ are random uniformly distributed samples in the volume Ω

According to the Law of Large Numbers, as long as the variance of $Y = f(X)$ is finite, the estmator is consistent in the sense that

$$\lim_{N\to\infty} \bar{Y} = \lim_{N\to\infty} \left[\frac{1}{N}\sum_k f(X_k)\right] = \bar{f}$$

where

$$\bar{f} = |\Omega|^{-1} \int_\Omega f(x)\, d^n x$$

The variance in the estmator is obtained from the Central Limit Theorem, which states that in the limit $\bar{Y}$ becomes Gaussian

The standard deviation or <u>error estimate</u> of $\hat{Y} \approx \overline{Y}$ is

$$\sigma(\hat{Y}) \approx \frac{1}{\sqrt{N}} \left[\int_\Omega \left(f(x) - \overline{f} \right)^2 dx \right]^{1/2}$$

$\frac{1}{\sqrt{N}}$: characteristic slow convergence of MC

This means that the MC estimate obtained for reasonable values of N will only be reasonably accurate if the integrand

$$f(x) = \varphi(x) \frac{e^{-\mathcal{H}(x)/k_B T}}{\mathcal{Z}} \approx \overline{f} = \text{const.}$$

But recall that in the thermodynamic limit (large systems), $e^{-\beta H(z)}$ becomes very localized around the most probable (equilibrium) states, meaning that only a very small fraction of the microstates contribute to the macro. observables.

Ideally, we would like to generate samples directly from the canonical distribution,

$$\rho(z) \sim \frac{e^{-\beta H(z)}}{\mathcal{Z}} = \rho_{GB}(z)$$

and then just average over them

$$Z^{-1}\int \varphi(z)\, e^{-\beta H(z)}\, dz \approx \frac{1}{N}\sum_{k=1}^{N} \varphi(z_k)$$

But how to generate samples from the Gibbs-Boltzmann distribution???

If we run <u>Hamiltonian dynamics</u> (MD) and just take snapshots of the system every so-often, the ergodic hypothesis says that we can obtain expectation values as simple averages over a long trajectory (or trajectories).

The problem is, of course, that the natural Hamiltonian dynamics will be too slow and spend most of its time "vibrating" around metastable or boring states; e.g.,

- vibrations around a crystal lattice site in between diffusive hops (e.g., vacancies in metals, spins in ferromagnets)
- ballistic motion of gas molecules in between collisions with other particles
- lots of thermal (non-reactive) collisions in between reactive collisions

Instead, we can just invent an artificial dynamics that jumps from interesting state to interesting state.

The invariant or equilibrium distribution of the dynamics must still be $\mathcal{P}_{GB}$ (7)

A simple way to get such dynamics is to construct a Markov Chain, that is, rules for the probability to jump from state S_i to state S_j (assumed discrete for notational simplicity) that are memory-free!

$$P(X_t = S_j \mid X_{t-1} = S_i, X_{t-2} = S_{i'}, \ldots) =$$

$$P(X_t = S_j \mid X_{t-1} = S_i) = P_{ij}$$

transition probability $P_{ij} = P(S_i \to S_j)$

$M = \{P_{ij}\}$ is the transition matrix (stochastic matrix)

$$\sum_j P_{ij} = 1 \quad , \quad P_{ij} \geq 0$$

Simulating a trajectory of such a chain is trivial: given present state i, choose one of the other states j with prob P_{ij}

The equilibrium distribution is a set of probabilities $\pi_i = P(X_t = S_i)$ that satisfy

BALANCE $\left[\pi_j = \sum_i \pi_i P_{ij} \;, \; \sum_i \pi_i = 1 \right.$

($\pi_i P_{ij}$: influx)

We want $\pi_i = \frac{e^{-\beta \mathcal{H}(S_i)}}{\mathcal{Z}}$

It is often better to make the chain reversible by satisfying the stronger detailed BALANCE condition $\rightarrow$ $\boxed{\pi_i P_{ij} = \pi_j P_{ji}}$

The detailed balance condition ensures that there is no flow cycles in phase space at equilibrium. This is based upon microscopic reversibility but is only valid at equilibrium (no dissipation). For example, chemical reactions that are irreversible cannot meet the detailed balance criterion.

Question:

How to construct a simple chain that obeys $\pi_i P_{ij} = \pi_j P_{ji}$

Answer: Metropolis Monte Carlo

is based on the idea of rejection:

Start with a naive symmetric chain

$$\tilde{P}_{ij} = \tilde{P}_{ji}$$

and then accept/reject trial moves with probability:

$$\begin{cases} \text{If } \dfrac{\pi_j}{\pi_i} \geq 1, \text{ accept } S_j \\ \text{otherwise, accept with probability } \dfrac{\pi_j}{\pi_i} \end{cases}$$

$$\boxed{P_{ij} = \tilde{P}_{ij} \min\left(1, \frac{\pi_j}{\pi_i}\right)}$$

Since $$\pi_i P_{ij} = \pi_i \tilde{P}_{ij} \min\left(1, \frac{\pi_j}{\pi_i}\right)$$

$$= \pi_j P_{ji} = \pi_j \tilde{P}_{ij} \min\left(1, \frac{\pi_i}{\pi_j}\right)$$

↑ recall symmetric

the Metropolis construction obeys detailed balance and thus the Gibbs distribution is the equilibrium distribution

Note $$\begin{cases} \frac{\pi_j}{\pi_i} = \exp\left[-\frac{\Delta H}{k_B T}\right] \\ \text{involves only the change in the energy and can be calculated in } O(1) \text{ operations in lots of cases.} \end{cases}$$

The field of equilibrium MC is all about coming up with better rules for $\tilde{P}_{ij}$: Minimize rejection and reduce correlation times.

Even though the Metropolis sampler the ideal importance sampling function, it leads to correlated samples which increases the variance (usually still $\sim \frac{1}{\sqrt{N}}$ but with bad prefactors that depend on system size).

We won't discuss this here, since we care about non-equilibrium!