Physics 262
Topic 3 (addendum): Numerical Simulations

Erel Levine
Nov 2015

7. MONTE-CARLO SIMULATIONS

7.1. Importance sampling

We want to calculate

$$\langle A\rangle = \sum_{\mathcal{C}} A(\mathcal{C})\frac{e^{-\beta E(\mathcal{C})}}{Z} \ .$$

The problem: we may not know how to calculate Z.

Enumeration will take forever (very many configurations).

Sampling: choose n configurations at random with probability $h(\mathcal{C})$, and estimate

$$\overline{A}_n = \frac{\sum_{i=1}^{n} A(\mathcal{C}_i)e^{-\beta E(\mathcal{C}_i)}/h(\mathcal{C}_i)}{\sum_{i=1}^{n} e^{-\beta E(\mathcal{C}_i)}/h(\mathcal{C}_i)} \ .$$

Why? because this gives an unbiased estimate:

$$\lim_{n\to\infty}\overline{A}_n = \frac{\sum_{\mathcal{C}} A(\mathcal{C})h(\mathcal{C})e^{-\beta E(\mathcal{C})}/h(\mathcal{C})}{\sum_{\mathcal{C}} e^{-\beta E(\mathcal{C})}h(\mathcal{C})/h(\mathcal{C})} = \frac{\sum_{\mathcal{C}} A(\mathcal{C})e^{-\beta E(\mathcal{C})}}{\sum_{\mathcal{C}} e^{-\beta E(\mathcal{C})}} = \langle A\rangle \ .$$

The easiest thing is to just pick configurations at random, or in other words $h(\mathcal{C}) = 1/\Omega$ where as usual Ω is the number of configurations. In this case

$$\overline{A}_n = \frac{\sum_{i=1}^{n} A(\mathcal{C}_i)e^{-\beta E(\mathcal{C}_i)}}{\sum_{i=1}^{n} e^{-\beta E(\mathcal{C}_i)}} \ .$$

The problem is that most of the terms we're going to have in our sum are zeros (many many configurations with high energy...).

The best option is $h(\mathcal{C}) = \alpha e^{-\beta E(\mathcal{C})}$. This would give

$$\overline{A}_n = \frac{\sum_i A(\mathcal{C}_i)/\alpha}{\sum 1/\alpha} = \frac{1}{n}\sum_i A(\mathcal{C}_i)$$

so all configurations contribute equally. But clearly $\alpha = 1/Z$ which we don't know how to calculate!

7.2. Markov Chain Monte Carlo

The idea is to build the sample in a sequence, one configuration at a time, in a way that would asymptotically lead to $h(\mathcal{C}) \to \alpha e^{-\beta E(\mathcal{C})}$. A MCMC is therefore completely defined by a transition matrix $W(\mathcal{C}' \to \mathcal{C})$ that sets the probability of sampling configuration $\mathcal{C}'$ after we sampled configuration $\mathcal{C}$.

The probability that our tth sample is configuration $\mathcal{C}$ is given by

$$P_t(\mathcal{C}) = \sum_{\mathcal{C}'} W(\mathcal{C}' \to \mathcal{C})P_{t-1}(\mathcal{C}') \ .$$

As the simulation continues, we expect things to settle down into a stationary state, such that the probability of drawing configuration $\mathcal{C}$ becomes i-independent. At this time

$$P_{\rm st}(\mathcal{C}) = \sum_{\mathcal{C}'} W(\mathcal{C}' \to \mathcal{C}) P_{\rm st}(\mathcal{C}') \ .$$

which can be rewritten as

$$\sum_{\mathcal{C}'} W(\mathcal{C} \to \mathcal{C}') P_{\rm st}(\mathcal{C}) = \sum_{\mathcal{C}'} W(\mathcal{C}' \to \mathcal{C}) P_{\rm st}(\mathcal{C}') \ .$$

To make sure that $P_{\rm st}(\mathcal{C})$ is the Boltzmann distribution we need, we will make sure that the transition matrix $W(\mathcal{C} \to \mathcal{C}')$ obeys the *Detailed Balance* condition, which states

$$P_{eq}(\mathcal{C}') W(\mathcal{C}' \to \mathcal{C}) = P_{eq}(\mathcal{C}) W(\mathcal{C} \to \mathcal{C}') \text{ for all } \mathcal{C}, \mathcal{C}'.$$

First, it is clear that if W obeys this condition, then P_{eq} obeys the stationarity condition above. If we also make sure that W is ergodic (that is, there exist a path of finite probability from every configuration to every configuration), then this stationary solution is also unique.

Detailed balance can therefore be viewed as a practical way of getting the asymptotic distribution of the MCMC procedure to be the distribution of our choice. All that's required is to make sure that

$$\frac{W(\mathcal{C}' \to \mathcal{C})}{W(\mathcal{C} \to \mathcal{C}')} = e^{-\beta(\mathcal{H}(\mathcal{C}) - \mathcal{H}(\mathcal{C}')} \ .$$

However, detailed balance is generally believed to be a property of an equilibrium system, stating that all microscopic processes in equilibrium are independently balanced. While detailed balance guarantees stationarity and increase of entropy, it is a significantly stronger requirement. The debate about its necessity is long and convoluted, and goes back to Maxwell and Boltzmann. However, it has strong empirical evidence, for example via Onsager's relations.

A successful MCMC algorithm therefore needs to guarantee two things: Ergodicity, and detailed balance with respect to the Bolzmann distribution. Here is a general scheme.

1. start with a random configuration $\mathcal{C}_0$.
2. given a configuration $\mathcal{C}_t$, we need to choose configuration $\mathcal{C}_{t+1}$:
   (a) draw a candidate configuration $\mathcal{C}'_t$ that is similar to $\mathcal{C}_t$ (in the Metropolis algorithm below the two configurations are different by a single spin flip).
   (b) calculate $\Delta E = E(\mathcal{C}'_t) - E(\mathcal{C}_t)$. This is a small calculation, since the change in energy is restricted to the local environment surrounding the change between the two configurations.
   (c) If $\Delta E < 0$ accept the candidate configuration, and let $C_{t+1} = C'_t$.
   (d) if $\Delta E > 0$ accept the candidate configuration with probability $e^{-\beta \Delta E}$. Otherwise reject it, and go back to try another candidate.
3. Repeat this until you have n configurations. Then calculate $\overline{A}_n = \frac{1}{n} \sum_t A(\mathcal{C}_t)$.

Indeed, if we take two configurations, $\mathcal{C}_1$ and $\mathcal{C}_2$ such that $E(\mathcal{C}_1) < E(\mathcal{C}_2)$ then

$$\frac{P_{st}(1)}{P_{st}(2)} = \frac{W(2 \to 1)}{W(1 \to 2)} = \frac{1}{e^{-\beta(E_2 - E_1)}} = \frac{e^{-\beta E_1}}{e^{-\beta E_2}}$$

as needed.

7.3. Metropolis algorithm for the Ising model:

This is a particular implementation of the MCMC idea, where two consecutive configurations differ by at most single spin. It it the easiest to implement. However, it is particularly bad at low temperatures, where most spins are surrounded by like-spins and the probability of actually making a move is very slow. In particular, it suffers a lot from critical slowing down, discussed below.

1. pick an initial configuration at random $\mathcal{C}_0 = \{s_i\}$.

2. main loop:

   (a) pick a site i at random.

   (b) calculate $\Delta E = E(-s_i, \{s_i\}) - E(s_i, \{s_i\})$.

   (c) if $\Delta E < 0$, $s_i(t+1) = -s_i$. Otherwise, set

$$s_i(t+1) = \begin{cases} -s_i(t) & \text{with probability } e^{-\beta \Delta E} \\ s_i(t) & \text{with probability } 1 - e^{-\beta \Delta E} \end{cases}$$

   (d) evey N such iterations are called a Monte-Carlo sweep. Record the configuration at the end of a sweep $\{s_i(t)\}$ and repeat.

3. When you're done collecting n configurations, calculate averages:

$$\overline{A} = \frac{1}{n} \sum_{sweeps} A[\{s_i(t)\}].$$

   For example,

$$\overline{m} = \frac{1}{n} \sum_{sweeps} \frac{1}{N} \sum_{i=1}^{N} s_i(t), \qquad \overline{U} = \frac{1}{n} \sum_{sweeps} E(\{s_i(t)\})$$

7.4. Critical slowing down

In a dynamical system, the dynamical variables are functions of time. An interesting quantity is the relaxation time, which is the time scale over which the system approaches equilibrium. If is a quantity which relaxes towards its equilibrium value , the the relaxation time can be defined as

$$\tau_A = \frac{\int dt\, t(A(t) - \langle A \rangle)}{\int dt (A(t) - \langle A \rangle)} .$$

Clearly if $(A(t) - \langle A \rangle) \sim e^{-t/T}$ then this definition would yield $\tau_A = T$ as expected. The relaxation time is intimately related with the correlation length. As the latter becomes larger, it takes more transitions to move from one configuration to an uncorrelated configuration. Near the critical temperature, the relaxation time of the order parameter m is related with the correlation length through $\tau \sim \xi^z$, where z is the dynamical critical exponent. The relaxation time therefore diverges near the critical temperature as

$$\tau \sim \xi^z \sim |T - T_c|^{-\nu z}.$$

This phenomenon is called critical slowing down, and poses a significant problem for computer simulations. Cluster-based algorithms (Swendsen-Wang, Wolf and others) are designed to address this issue.