Mechanics – Lecture 14 – 5/3/2022

[On 5/3 we'll do some of the material in the "Lecture 13" notes, plus what's here.]

How can we sample the canonical ensemble numerically?

I'll focus on systems with finitely many states. (Lots of interesting examples have this character, eg the 2D Ising model.) Widely used method: Markov chain Monte Carlo. I'll focus on one basic (but widely used) example: Metropolis sampling.

Basic idea: design a Markov chain st its (unique) equil. distrn is the one you want to sample.

Background on Markov chains (on a finite state space, let's label the states $1, \ldots, n$)

Let p_{ij} = prob of transition from i to j

(warning: in some books, incl Buhler's, this is P_{ji}). Clearly

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

For the assoc random walk, let

$\rho(i,t)$ = prob of being at state i at time t.

Then

$$\rho(i,t+1) = \sum_j \rho(j,t)\, p_{ji}$$

So a distrn is <u>stationary</u> if

$$\pi_i = \sum_j{}' \pi_j p_{ji}$$

So, basic idea is: given $\{\pi_i\}$ (eg the distrn assoc canonical ensemble), we want to find p_{ij} st $\{\pi_i\}$ is the stationary distrn +

$$\rho(i,n) \to \pi_i \text{ as } n \to \infty.$$

<u>Sufficient conds for this</u>: both

(a) $\pi_j p_{ji} = \pi_i p_{ij}$ for each $i+j$ ("detailed balance")

(b) The chain is irreducible (for any $i+j$, prob of starting at i + getting to j in N steps is pos. for some $N>0$; equivalently: $(p^N)_{ij} > 0$ for some N.))

hold.

Intuition about (a): it says "prob flux from j to i = prob flux from i to j" in equilibr.

Note that (a) <u>implies</u> that π is stationary, since

$$\sum_j{}' \pi_j p_{ji} \underset{\text{detailed balance}}{=} \sum_j{}' \pi_i p_{ij} \underset{\sum_j{}' p_{ij} = 1}{=} \pi_i$$

(However, detailed balance is not in general equivalent to π being stationary.)

This argument gives a sense why detailed balance is useful: let

$$D_t = \sum_i{}' |\rho(i,t) - \pi_i|$$

Then detailed balance assures at least

$$D_{t+1} \leq D_t \,.$$

Proof:

$$\begin{aligned} D_{t+1} &= \sum_i{}' |\rho(i,t+1) - \pi_i| \\ &= \sum_i{}' \Big| \sum_j{}' \rho(j,t) p_{ji} - \pi_i p_{ij} \Big| \\ &= \sum_i{}' \Big| \sum_j{}' \rho(j,t) p_{ji} - \pi_j p_{ji} \Big| \\ &\leq \sum_i{}' \Big(\sum_j{}' |\rho(j,t) - \pi_j| \Big) p_{ji} \\ &= D_t \end{aligned}$$

[One must work harder to show $D \to 0$ as $t \to \infty$; in particular, irreducibility has to be used.]

OK, but how to choose p_{ij}? Metropolis sampling works as follows:

step 1: Start with any irreducible, symmetric Markov chain ("symmetric" means $p_{ij} = p_{ji}$). For example: in Ising model one step of chain could involve flipping one randomly chosen spin.

Step 2: construct new Markov chain as follows:

$$p^*_{ij} = \begin{cases} p_{ij} & \text{if } \pi_j > \pi_i \\ p_{ij}\dfrac{\pi_j}{\pi_i} & \text{if } \pi_j < \pi_i \end{cases} \quad \text{if } i \neq j$$

$$p^*_{ii} = p_{ii} + \sum_{\pi_j < \pi_i}' p_{ij}\,(1 - \pi_j/\pi_i)$$

[Operationally: when original random walker would go from i ~~to~~ j st $\pi_j > \pi_i$, so does the new one. But when orig walker would go from i to j st $\pi_j < \pi_i$, accept such moves only sometimes (with prob π_j/π_i); else stay at i.]

It's easy to see that $\pi_j p^*_{ji} = \pi_i p^*_{ij}$, since

$\pi_j < \pi_i \Rightarrow$ LHS is $\pi_j p_{ji}$

RHS is $\pi_i \dfrac{\pi_j}{\pi_i} p_{ij} \underset{\text{symmetry}}{=} \pi_j p_{ji}$

(the case $\pi_j > \pi_i$ is similar).

<u>A key pt</u>: we only need π_j/π_i, not π_i + π_j individually. This is important! In canonical ensemble,

$$\pi_j = \frac{1}{Z_\beta} e^{-\beta H_j}$$

and finding $Z_\beta = \sum_j e^{-\beta H_j}$ requires evaluating a huge sum. But

$$\frac{\pi_j}{\pi_i} = e^{-\beta(H_j - H_i)}$$

is easy to evaluate

As indicated earlier at step 2, implementation is easy: to take a step of the walk assoc p^*,

<u>first</u> take a step of the walk assoc to p (eg for Ising, flip one spin)

<u>Then</u> consider whether proposed new state j has larger prob ($\pi_j > \pi_i$). If yes, accept it. If not, accept it with prob π_j/π_i. (If move is rejected, stay at state i.)

<u>An awkward pt</u>: to sample the distrn assoc to π you really want many independent samples. To get this you can run the Markov chain a long time + sample once in a while. How often depends on how long it takes for correlations to decay. (For the naive approach to Ising this takes <u>many</u> moves.)