<u>Mechanics</u> - <u>Lecture 12</u> - <u>5/1/2019</u>

[Actually, much of the 5/1 lecture will be spent finishing the Lecture 11 notes; so the material here will certainly spill over to 5/8]

These notes: <u>a little more discussion of the canonical distrn</u>; then <u>Metropolis sampling</u> (as a way to sample this distrn)

<u>Q</u>: What is the prob that $H \in (E, E+dE)$ under the canonical distrn?

<u>Ans</u>:

in our discrete setting,

$$\frac{1}{Z(\beta)} e^{-\beta E} \Omega(E) = \frac{1}{Z(\beta)} e^{[S(E)-\beta E]}$$

writing (as usual) $S(E) = \ln \Omega(E)$

in a cont's setting

$$\frac{1}{Z(\beta)} \int_{H=E} \frac{1}{|\nabla H|} e^{-\beta H} d(\text{area})$$

Focusing on the discrete setting, we see that the most likely "energy" E is assoc to the min value of

$$\text{"free energy"} = E - \frac{1}{\beta} S(E)$$

$= E - TS$

Intuition: a value $H = E$ becomes more likely either because it's lower ($\Rightarrow$ higher density $e^{-\beta H}$) or because it's achievable many different ways (larger entropy). At finite temperature, the most likely value of H is determined by min of free energy (not min of H itself)

Q: Can we ever do by-hand calculations with the canonical distn?

Well, sometimes (but rarely). Bühler offers some examples in Sec 3.4. Here's just one: a particle in a 3D box $[0,L]^3$, otherwise free (reflecting at the bdry)

phase space: (x, y, z, p_x, p_y, p_z)

Hamiltonian: $\frac{1}{2}(p_x^2 + p_y^2 + p_z^2)$ if $m=1$

$\Rightarrow$ normalization constant is

$$Z(\beta) = \int e^{-\beta H} = L^3 \int_{\mathbb{R}^3} e^{-\beta|\xi|^2/2} d\xi = L^3 \cdot \left(\frac{2\pi}{\beta}\right)^{3/2}$$

Avg energy is

$$\langle H \rangle = -\frac{\partial}{\partial\beta}\ln(Z) = \frac{3}{2\beta} = \frac{3}{2}T$$

Note: calcn is feasible because H is quadratic (so integral involves a Gaussian).

How can we sample the canonical distrn numerically? I'll focus on systems with finitely many states for this discn (lots of interesting examples have this character, eg 2D Ising described at end of these notes). Widely used method: "Markov chain Monte Carlo." I'll discuss one basic (but widely used) example: "Metropolis sampling".

Basic idea: design a Markov chain st its (unique) equil distr is the canonical distrn.

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

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

(warning: in Butler's book this is P_{ji}). Observe that

$$p_{ij} \geq 0 \; ; \quad \sum_j' p_{ij} = 1$$

For the random walk assoc to this process, let $\rho(i,n)$ = prob of being at state i at t'e n. Then

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

so a prob. distrn π_i is stationary if

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

Our idea is: given $\{\pi_i\}$ (the canonical distrn, in our case) let's find p_{ij} st $\{\pi_i\}$ is the stat distrn, + st. $\rho(i,n) \to \pi_i$ as $n \to \infty$.

Sufft conds: 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 startg at i + getting to j in n steps is >0 for some n; equiv: $(p^n)_{ij} > 0$ for some n

Interpretn of (a): for equil distr π, "prob. flux from j to i" = "prob flux from i to j". This implies that π is stationary for p, since

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

but it is not in general equiv to π being stationary for p.

Arg't to explain why (a) is useful: let

$$D_n = \sum_i{}' |\rho(i,n) - \pi_i|$$

Then

$$D_{n+1} = \sum_i |\rho(i,n+1) - \pi_i|$$

$$= \sum_i{}' \left| \sum_j{}' \rho(j,n) p_{ji} - \pi_i p_{ij} \right|$$

$$= \sum_i{}' \left| \sum_j{}' \rho(j,n) p_{ji} - \pi_j p_{ji} \right|$$

$$\leq \sum_i{}' \left(\sum_j{}' |\rho(j,n) - \pi_j| \right) p_{ji}$$

$$= D_n$$

[This shows D_n is nondecreasing; to show convergence to 0 one must work more — eg using irreducibility]

OK, how to choose p_{ij}?

<u>step 1</u> start with <u>any</u> irreducible, symmetric Markov chain ("symmetric" means: $p_{ij} = p_{ji}$)
For example: in an Ising model one step of the chain might consist of flipping one randomly chosen spin.

<u>step 2</u> construct a <u>new</u> Markov chain as follows:
The new transition prob p^*_{ij} is

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

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

It's easy to see that $\pi_j p^*_{ji} = \pi_i p^*_{ij}$, so detailed balance holds.

<u>Key pts</u>:

① we only need π_j/π_i, not π_j & π_i individually. So for Canonical distrn, we need only $\Delta H = H(X_i) - H(X_j)$ (no need for the partition fn $Z(\beta)$, which is usually inaccessible)

② Implementation is easy. To take one step of the random walk assoc to p^*, starting from state i

first take a step of the process assoc to p_{ij} (eg for Ising, flip a randomly chosen spin)

Then consider the proposed new state j. If it is more likely ($\pi_j > \pi_i$) accept it. If it is less likely ($\pi_j < \pi_i$), accept it with prob π_j/π_i. If j is rejected, stay at i.

Regardless of initial state, if N is large, distrn of values of $\rho(i,N)$ is approx π_i.

2D "Ising model" would be an example to explore this way. It involves an $N \times N$ lattice of "spins" (in 2D case) $s_{ij} = \pm 1$,

$$H = -\tfrac{1}{2} \sum' s_{ij} \left(s_{i+1,j} + s_{i-1,j} + s_{i,j+1} + s_{i,j-1} \right)$$

(using periodic bc say) so spins fully aligned is the lowest-energy state (but very unlikely). Most interesting observable: $\mu = \frac{1}{N^2} \sum' s_{ij}$ ("magnetization") Fact: at low temp (large β), the likely values of

μ (in Canonical distr^n) cluster near $\pm\mu_*$ (not zero!). When temp is high (low β) the likely values cluster instead near 0. This is an example of a "phase transition".