

Mechanics - Lecture 12 - 4/25/2018 (pg 12.3 corrected 5/2)

These notes:

- further descri of the canonical distro
- Metropolis sampling (a way to sample this distro)

Q: What is the prob that  $H \in (E, E+dE)$  under the canonical distro?

Ans / in a 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,

$$\frac{1}{Z(\beta)} \int_{H=E} \frac{1}{|\partial H|} e^{-\beta H} \mathcal{Q}(\text{Area})$$

Focusing on the discrete setting, we see that the most likely "energy"  $E$  is assoc to 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 dist'n?

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

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

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

Avg energy is

$$\langle H \rangle = - \frac{\partial \ln(Z)}{\partial \beta} = \frac{S}{Z\beta} = \frac{S}{ZT}$$

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

How can we sample the canonical distn numerically? I'll focus on systems with finitely many states for this descr (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 distn is the canonical distn.

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

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

(Warning: in Bulter'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  $f(i, n) =$  prob. of being at state  $i$  at time  $n$ . Then

$$f(i, n+1) = \sum_j f(j, n) P_{ji}$$

so a prob. distribn  $\pi_i$  is stationary if

$$\pi_i = \sum_j \pi_j P_{ji}$$

Our idea is: given  $\{\pi_i\}$  (the canonical distribn, in our case) let's find  $P_{ij}$  st.  $\{\pi_i\}$  is the stat. distribn, + st.  $f(i, n) \rightarrow \pi_i$  as  $n \rightarrow \infty$ .

Sufft cond: 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  $> 0$  for some  $n$ ; equiv:  $(P^n)_{ij} > 0$  for some  $n$ )

Interpret of (a): at equil state  $i$ , "prob. flux from  $j$  to  $i$ " = "prob flux from  $i$  to  $j$ ". This implies that  $\pi$  is

stationary for  $p$ , since

$$\sum_j \pi_j P_{ji} = \sum_j \pi_i P_{ij} = \pi_i$$

↑ detailed balance
 ↑  $\sum_j p_{ij} = 1$

but it is not in general equiv to  $\pi$  being stationary for  $p$ .

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

$$D_n = \sum_i |p(i, n) - \pi_i|$$

Then

$$\begin{aligned} D_{n+1} &= \sum_i |p(i, n+1) - \pi_i| \\ &= \sum_i \left| \sum_j p(j, n) P_{ji} - \pi_i P_{ii} \right| \\ &= \sum_i \left| \sum_j p(j, n) P_{ji} - \pi_j P_{ji} \right| \\ &\leq \sum_i \left( \sum_j |p(j, n) - \pi_j| \right) P_{ji} \\ &= D_n \end{aligned}$$

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

OK, how to choose  $p_{ij}^*$ ?

step 1 start with any 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.

step 2 construct a new Markov chain as follows:  
the new transition prob  $p_{ij}^*$  is

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

$$p_{ii}^* = p_{ii} + \sum_{\substack{j \\ \pi_j < \pi_i}} p_{ij} (1 - \pi_j / \pi_i)$$

It's easy to see that  $\pi_j p_{ji}^* = \pi_i p_{ij}^* > 0$   
detailed balance holds.

Key pts:

- we only need  $\pi_j / \pi_i$ , not  $\pi_j + \pi_i$  individually. So in canonical distns, 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. In 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  $\pi_j / \pi_i$ .  
 If  $j$  is rejected, stay at  $i$ .

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

2D "Ising model" would be an example to explore this way. It involves an  $N \times N$  lattice of "spins" (in 2D case)

$$s_{ij} = \pm 1,$$

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

(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  $\beta$ ), the likely values of

$\mu$  (in Canonical distn) cluster near  $\pm \mu_x$   
(not zero!). When temp is high (low  $\beta$ ) the  
likely values cluster instead near 0. This is  
an example of a "phase transition".