Edited Emails: Statistical Mechanics

Ying Nian Wu
UCLA Department of Statistics
ywu@stat.ucla.edu

Uniform distribution of isolated system

Consider an isolated big system X (e.g., the universe) whose energy E(X) is fixed at E. X evolves deterministically within the state space

Omega(E) = {X: E(X) = E}.

Assume that with enough time, X will visit every state in Omega(E) with equal frequency, an assumption called ergodicity, then at a random time, X follows a uniform distribution over Omega(E).

Gibbs distribution of a subsystem

Now consider a small subsystem x (e.g., a molecule), which is part of X, and let Y be the rest of X, i.e., X = (x, Y), and

E(X) = E(x) + E(Y) = E.

Then under the uniform distribution of X, the distribution of x is

p(x) = 1/Z(T) exp(- E(x)/T),

where T is the temperature. This is the origin of the Gibbs distribution, or Boltzmann distribution, or energy-based model.

The reason is as follows. For a given value x of the small subsystem, let y be the rest of the large system X, then

E(Y) = E(X) - E(x) = E - E(x).

Among all the states in Omega(E) = {X: E(X) = E}, the number of states where the subsystem is x is the same as the number of states in Omega1(E-E(x)) = {Y: E(Y) = E - E(x)}. Then

p(x) = |Omega1(E-E(x))| / |Omega(E)|.

Thus

log p(x) = log |Omega1(E-E(x))| + const1
         = - beta E(x) + const2

using the first-order Taylor expansion, where

beta = d/dE log |Omega1(E)|

is the derivative term in the first-order Taylor expansion.

We call beta = 1/T. Then

log p(x) = - E(x)/T + const2,

and

p(x) = 1/Z(T) exp(- E(x)/T).

We can interpret p(x) to be the distribution of a system exchanging heat with a big environment at temperature T, i.e., a heat bath.

Micro-canonical ensemble and canonical ensemble

Omega(E) = {X: E(X) = E} or the uniform distribution over Omega(E) is called micro-canonical ensemble, and p(x) is called canonical ensemble. Under p(x), the energy E(x) fluctuates, because x exchanges heat with its environment y. But if x is large enough with a large number of degrees of freedom, the average energy per degree of freedom converges to a constant due to law of large number or concentration of measure, and p(x) behaves like a micro-canonical ensemble.

Heat and entropy

For micro-canonical ensemble (or a large canonical ensemble),

S = log |Omega(E)|

is called entropy.

Since

beta = 1/T = dS / dE, or dS = beta dE,

the change of energy causes the change of entropy. The change of energy is in the form of heat. So if we inject heat into the system, its entropy will increase. If the system releases heat, its entropy will decrease.

Free energy

Suppose we want to extract the energy of a system x for work. Let us assume x is large enough so that we can approximate it by a micro-canonical ensemble. We hope to extract all the energy E(x) = e, to make it to zero. However, at E(x) = 0, the ensemble becomes {x: E(x) = 0}, and the entropy is also reduced to a minimum S0. For the system x and its environment y, the total

entropy should not decrease. Thus if the system reduces its entropy, the environment must increase its entropy. That is, the system must release heat to the environment. It is like you are tidying up your living room to make it organized, and in doing so, you reduce the entropy of your living room. However, you generate heat to the environment to increase the entropy of the environment.

At a fixed temperature, the change of entropy S - S0 is thus transferred to its environment in the form of heat T (S-S0). Therefore, we can only extract work

e - T(S-S0).

Since S0 is a constant, we can define e - TS as free energy.

When extracting work, it involves the change of volume V of the piston of the steam engine under a certain pressure p. Thermodynamics is mainly to study how to convert heat to work. In this process, we have to consider the change of entropy, i.e., the count of configurations in the ensemble.

Free energy = - T log Z

For canonical ensemble p(x), its energy $E(x)$ = energy(x) fluctuates, so we define free energy as the expectation of energy – T entropy, i.e.,

$$E\_p[energy(x)] – T \text{ entropy}(p) = E\_p[energy(x) + T \log p(x)]$$
$$= E\_p[energy(x) + T ((-energy(x)/T) - \log Z)]$$
$$= - T \log Z.$$

Variational approximation

For a normalizing flow model q(x),

$$KL(q|p) = E\_q[\log q] – E\_q[-energy(x)/T – \log Z]$$
$$= - \text{entropy}(q) + E\_q[energy(x)/T] + \log Z.$$

Thus

$$-T \log Z = - T \text{ entropy}(q) + E\_q[energy(x)] – KL(q|p).$$

So the free energy is upper-bounded by

$$F(q) = E\_p[energy(x)] – T \text{ entropy}(q).$$

We can minimize F(q) to obtain q.

Reinforcement learning

A related problem arises in reinforcement learning, where we want to find policy $q(x)$ (x now becomes action, and we make state implicit):

$F(q) = E\_q[cost(x)] + lambda\ KL(q|p0),$

where p0 is a base policy.

The optimal q is

$p(x) = 1/Z\ exp(-\ beta\ cost(x))\ p0(x),$

with beta = 1/lambda, because

$KL(q|p) = E\_q[log\ q + beta\ cost(x) - log\ p0 + log\ Z]$
$\qquad = KL(q|p0) + beta\ E\_q[cost(x)] + log\ Z$
$\qquad = F(q)/lambda + log\ Z,$

which is minimized at q = p.