Lecture 2: Meanfield approximation, variational principle and Landau expansion

Naoki KAWASHIMA[1]

[1]ISSP, U. Tokyo

April 15, 2019

[2-1] Mean-field approximation

In this lecture we will see:

- Molecular-field theory does not gives us the free energy.

- Gibbs-Bogoliubov-Feynman inequality give us a very systematic and frexible framework for constructiong the mean-field-type approximations.

Molecular field theory revisited

- In the molecular-field theory, the effect of **environment** is replaced by an additional term, in the case of Ising ferromagnet, we focus on a single spin, say S_0, and replace the Hamiltonian as

$$\mathcal{H} = -J\sum_{ij} S_i S_j - H\sum_i S_i \; \rightarrow \; \mathcal{H}_{\textbf{MF}} = -H_{\mathrm{MF}}S_0 - HS_0$$

- It is also argued that the right choice of the effective field is

$$H_{\mathrm{MF}} = J\sum_j \langle S_j \rangle$$

- The uniformity condition $m = \langle S_i \rangle$ (independent of i) yields,

$$m = \tanh(\beta(H + zJm)) \quad (z = \text{(number of nearest-neighbors)})$$

Why should we complain?

- In principle, we have multiple solutions of the self-consistent equation.

- The molecular-field theory does not tell us which solution we should choose.

- If the theory allowed us to compute the free energy for each solution, we would be able to tell which one to take.

Gibbs-Bogoliubov-Feynman (GBF) inequality

Theorem 1 (GBF inequality)

For two Hamiltonians $\mathcal{H}(S)$ and $\mathcal{H}_0(S)$ defined on the same space $S \in \Omega$,

$$F_{\mathrm{v}} \equiv F_0 + \langle \mathcal{H} - \mathcal{H}_0 \rangle_0 \geq F, \qquad (1)$$

where F and F_0 are the free-energies of $\mathcal{H}$ and $\mathcal{H}_0$ respectively and $\langle \cdots \rangle_0$ is the thermal average with respect to $\mathcal{H}_0$.

Variational calculation

When $\mathcal{H}(S)$ is the Hamiltonian of the system that we want to study but is not solvable, by taking $\mathcal{H}_0(\mathbf{\Lambda}, S)$ for $\mathcal{H}_0$ in (1) that also depends on a list of parameters $\mathbf{\Lambda}$ and is solvable for any $\mathbf{\Lambda}$, $F_v(\mathbf{\Lambda})$ gives us the upper bound of the correct free energy. (And it is computable!)

GBF inequality from information-scientific view-point

Theorem 2 (Relation to Kullback-Leibler divergence)

The "error" in the variational free-energy is proportional to the Kullback-Leibler divergence of the thermodynamic distribution of $\mathcal{H}_0$ relative to that of $\mathcal{H}$.

More precisely,

$$F_{\mathrm{v}} - F = k_{\mathrm{B}} T \; I_{\mathrm{KL}}[\rho_0 | \rho] \tag{2}$$

where

$$\rho_0 \equiv e^{-\beta \mathcal{H}_0}/Z_0, \quad \text{and} \quad \rho \equiv e^{-\beta \mathcal{H}}/Z \tag{3}$$

and

$$I_{\mathrm{KL}}[P|Q] \equiv \sum_S P(S) \log \frac{P(S)}{Q(S)} \tag{4}$$

(Z and Z_0 are partition functions of $\mathcal{H}$ and $\mathcal{H}_0$ respectively.)

Proof of Theorem 2

$$I_{\mathrm{KL}} \left[Z_0^{-1} e^{-\beta \mathcal{H}_0} \,\middle|\, Z^{-1} e^{-\beta \mathcal{H}} \right]$$

$$= \sum_S \left(\frac{e^{-\beta \mathcal{H}_0(S)}}{Z_0} \log \left(\frac{e^{-\beta \mathcal{H}_0(S)}}{Z_0} \right) - \frac{e^{-\beta \mathcal{H}_0(S)}}{Z_0} \log \left(\frac{e^{-\beta \mathcal{H}(S)}}{Z} \right) \right)$$

$$= -\log Z_0 + \langle -\beta \mathcal{H}_0 \rangle_0 + \log Z - \langle -\beta \mathcal{H} \rangle_0$$

$$= \beta F_0 - \beta \langle \mathcal{H}_0 \rangle_0 - \beta F + \beta \langle \mathcal{H} \rangle_0$$

$$= \beta (F_{\mathrm{v}} - F)$$

Kullback-Leibler information measure is positive

$$I_{\mathrm{KL}}[P|Q]$$
$$= \sum_S P(S) \log \frac{P(S)}{Q(S)} = -\sum_S P(S) \log \frac{Q(S)}{P(S)}$$
$$\geq -\sum_S P(S) \left(\frac{Q(S)}{P(S)} - 1 \right) \quad (\text{because } \log(x) \leq x - 1)$$
$$= \sum_S (Q(S) - P(S)) = 1 - 1 = 0$$

Remark

This inequality together with Theorem 2 proves Theorem 1.

Quantum extention

- Though the theorems have been proved for classical systems, the corresponding quantum version of them can be also proved.
- For the extension, the KL divergence must be generaized to

$$I_{\mathrm{KL}}[P|Q] \equiv \mathrm{Tr}\left(P(\log P - \log Q)\right) \tag{5}$$

  where P and Q are now density operators satisfying $\mathrm{Tr}\left(P\right) = \mathrm{Tr}\left(Q\right) = 1$.

- Only non-trivial part in the proof of the quantum extension is the positivity of the KL information. The rest is straight-forward simply by replacing $\sum_S$ by Tr.

Proof of quantum extention (1)

Theorem 3 (Positivity of quantum KL divergence)

For any density operators P and Q,

$$I_{\mathrm{KL}}[P|Q] \equiv \mathrm{Tr}\left(P \log P - P \log Q\right) \geq 0 \tag{6}$$

Proof:

- Let us take the basis set in which P is diagonal, i.e., $P_{ij} = p_i \delta_{ij}$.
- For some unitary operator U, $Q_{ij} = u_{ik} q_k u_{jk}^*$
- With this u_{ij},

$$I_{KL}[P|Q] = \sum_i \left\{ p_i \log p_i - p_i \sum_j a_{ij} \log q_j \right\} \quad (a_{ij} \equiv |u_{ij}|^2)$$

Proof of quantum extention (2)

- Now, notice that $a_{ij} \geq 0, \quad \sum_i a_{ij} = \sum_j a_{ij} = 1$.
- Let us define $p'_{ij} = p_i a_{ij}, \quad q'_{ij} = q_j a_{ij}$.
- Then, these can be regarded as the classical distribution function in the squared Hilbert space $H \times H \equiv \{(ij) \,|\, i, j \in H\}$:

$$p'_{ij} \geq 0, \ q'_{ij} \geq 0, \ \sum_{ij} p'_{ij} = \sum_{ij} q'_{ij} = 1$$

- Now, we can see

$$\sum_{ij} p'_{ij} \log p'_{ij} = \sum_i p_i \log p_i + \sum_{ij} p_i a_{ij} \log a_{ij}$$

$$\sum_{ij} p'_{ij} \log q'_{ij} = \sum_{ij} p_i a_{ij} \log q_j + \sum_{ij} p_i a_{ij} \log a_{ij},$$

- Thus we have obtained $I_{KL}[P|Q] = \sum_{ij} p'_{ij} \log(p'_{ij}/q'_{ij}) \geq 0$
  (because the RHS is the classical KL information).   (QED)

Variational approximation to the Ising model (1)

- For the target Hamiltonian $\mathcal{H} \equiv -J \sum_{ij} S_i S_j - H \sum_i S_i$, let us take the "trial" Hamiltonian $\mathcal{H}_0 \equiv -\Lambda \sum_i S_i$. where Λ is a variational parameter.

- Then, our variational free energy is

$$F_v = F_0 + \langle \mathcal{H} - \mathcal{H}_0 \rangle_0$$
$$= \langle \mathcal{H} \rangle_0 - S_0 T$$

  where $S_0 \equiv T^{-1}(\langle \mathcal{H}_0 \rangle_0 - F_0)$ is the entropy of the $\mathcal{H}_0$ system

- By introducing $m \equiv \langle S_i \rangle_0 = \tanh \beta \Lambda$,

$$\langle \mathcal{H}_0 \rangle_0 = -\frac{z}{2} N J m^2 - H N m \qquad S_0 = N \sigma(m) \tag{7}$$

$$\sigma(m) \equiv -k_B \left(\frac{1+m}{2} \log \frac{1+m}{2} + \frac{1-m}{2} \log \frac{1-m}{2} \right) \tag{8}$$

Variational approximation to the Ising model (2)

Variational free-energy density

$$f_{\mathrm{v}} \equiv \frac{F_{\mathrm{v}}}{N} = -\frac{zJm^2}{2} - Hm - T\sigma(m)$$

- The GBF inequality tells us that we should minimize f_{v} with respect to λ. Since f_{v} depends on λ only through m, the stational condition $\partial f_{\mathrm{v}}/\partial \lambda = 0$ leads to $\partial f_{\mathrm{v}}/\partial m = 0$.

- From this, we obtain the same as the molecular-field approx.:

$$\frac{\partial f_{\mathrm{v}}}{\partial m} = 0 \quad \Rightarrow \quad m = \tanh \beta H_{\mathrm{MF}} \quad (H_{\mathrm{MF}} \equiv zJm + H). \tag{9}$$

Mean-field free energy — Landau expansion

- We have also obtained the explicit expression for the free energy.
- Since its behavior near $m \approx 0$ is most important for the critical phenomena, let us expand f_v with respect to m.

$$f_v = -\frac{zJ}{2}m^2 - Hm - k_B T \left(\log 2 - \frac{m^2}{2} - \frac{m^4}{12} \right) \tag{10}$$

- From the condition (coefficient of m^2) $= 0$, we obtain $k_B T_c = zJ$.
- Near $T \approx T_c$, we can get the Landau expansion:

$$f_v \approx f_v^0 + tm^2 + um^4 - Hm \tag{11}$$

where $f_v^0 \equiv -zJ \log 2$, $t \equiv (k_B T - zJ)/2$ and $u \equiv zJ/12$

[2-2] Summary of Lecture 2

- Molecular-field theory does not give us the free energy.

- Gibbs-Bogoliubov-Feynman inequality give us a very systematic and frexible framework for constructiong the mean-field-type approximations that also provides the free energy.

- Landau expansion is useful in having a clear view of the phase transitions.

Homework

- Veryfy (7) and (8).

- At the point where we have arrived at (9), m is just a variational parameter and its physical meaning is not yet clear. Give the reason why we can interpret it as the magnetization.

- Following the example of the Ising model in the lecture, obtain the Landau expansion of the 3-state Potts model by taking the trial Hamiltonian

$$\mathcal{H}_0 \equiv -\Lambda \sum_i \delta_{S_i,1}. \tag{12}$$

  This time, the order parameter should be $m \equiv \langle \delta_{S_i,1} \rangle - 1/3$. What is the essential difference from the Ising case?

- Submit your report on one of these problems at the beginning of the next lecture.