Mean Field Asymptotics in Statistical Learning.

Jan 25 th.

Chapter 2
[MM 08]

Lecture 2 : Basics concepts in statistical physics.

Thermodynamics : heat temperature work.

Macroscopic. 4 laws

Statistical physics : microscopic.

Static dynamical

Most important message: The free energy function is most important quantity to calculate in SP. It can be used to derive most of the properties of physical system.

① Configuration space and Gibbs distribution.

Concept.

Configuration space : Ω

configuration : $\sigma \in \Omega$

Observable : $f : \Omega \to \mathbb{R}$

Energ function : $H : \Omega \to \mathbb{R}$
(Hamiltonian)

Example (Ising model).

$\Omega = \{\pm 1\}^n$ $\quad (\Omega = \mathbb{R}^n)$

$\sigma = \{+1, -1, +1, +1, -1, \cdots\}$

$f(\sigma) = \sum_{i=1}^{n} \sigma_i$

$H(\sigma) = \sum_{\substack{i,j \in [n] \\ i \neq j}} J_{ij} \sigma_i \sigma_j$

Ferromagnetic model : $J_{ij} = -1$

Spin glass model : $J_{ij} \overset{iid}{\sim} N(0, \sigma^2)$

☆ Gibbs / Boltzman distribution.

ν_0 is reference measure on Ω.

$$P_\beta(d\sigma) = \frac{1}{Z(\beta)} \exp\{-\beta H(\sigma)\} \nu_0(d\sigma).$$

$$Z(\beta) = \int_\Omega \exp\{-\beta H(\sigma)\} \nu_0(d\sigma)$$ Partition function.

β : inverse temperature. $\quad T = 1/\beta$: temperature.

Ensemble average : $f : \Omega \to \mathbb{R}$. $\quad \langle \cdot \rangle_\beta$

$$\langle f \rangle_\beta \equiv \int_\Omega f(\sigma) P_\beta(d\sigma),$$

High temperature limit : $\beta \to 0$, $T \to \infty$ $\quad P_\beta \to \nu_0$. $(\beta \to 0)$

low temperature limit : $\beta \to \infty$, $T \to 0$.

P_β concentrate on $\Omega_0 \equiv \underset{\sigma}{\operatorname{argmin}} H(\sigma)$

$$\lim_{\beta \to \infty} P_\beta(\Omega_0) = 1.$$

② Thermodyanamic potentials. (Special functions of β).

Fix $H = H_n$ $\quad P_\beta$

Free energy : $F(\beta) = -\frac{1}{\beta} \log Z(\beta)$.

Free entropy : $\Phi(\beta) = \log Z(\beta)$.

Internal energy : $U(\beta) = \langle H \rangle_\beta = \int_\Omega H(\sigma) P_\beta(d\sigma)$.

Canonical entropy : $S(\beta) = -\sum_{\sigma \in \Omega} P_\beta(\sigma) \log P_\beta(\sigma)$

a) $$\Phi'(\beta) = \frac{d}{d\beta}\left[\log Z(\beta)\right] = \frac{Z'(\beta)}{Z(\beta)} = -\langle H \rangle_\beta = -U(\beta).$$

$$Z'(\beta) = \frac{d}{d\beta} \int \exp\{-\beta H(\sigma)\} \nu_0(d\sigma)$$
$$= \int \exp(-\beta H(\sigma)) \times (-H(\sigma)) \nu_0(d\sigma).$$

b). $\Phi''(\beta) = \mathrm{Var}_\beta(H) = [\langle H^2 \rangle_\beta - \langle H \rangle_\beta^2]$.

$$\Phi''(\beta) = \frac{d}{d\beta}\left[\frac{Z'(\beta)}{Z(\beta)}\right] = \frac{Z''(\beta)Z(\beta) - Z'(\beta)^2}{Z(\beta)^2}$$

$$Z''(\beta) = \frac{d}{d\beta}\int \exp(-\beta H(\sigma)) \times (-H(\sigma))\, \nu_0(d\sigma)$$
$$= \int \exp(-\beta H(\sigma)) (H(\sigma))^2\, \nu_0(d\sigma).$$

$$\Phi''(\beta) = \langle H^2 \rangle_\beta - \langle H \rangle_\beta^2 \geq 0.$$

c) $S(\beta) = -\sum P_\beta(\sigma) \log P_\beta(\sigma)$

$= -\sum P_\beta(\sigma)\, [-\beta H(\sigma) - \log Z(\beta)]$

$= \beta \langle H \rangle_\beta + \log Z(\beta)$

$= \beta \cdot U(\beta) + \Phi(\beta)$

$$F(\beta) = -\frac{1}{\beta}\Phi(\beta) = U(\beta) - \frac{1}{\beta}S(\beta)$$

③ Thermodynamic limit and phase transition.

H_n

$F_n(\beta)$, $\Phi_n(\beta)$, $U_n(\beta)$, $S_n(\beta)$.

For large n, the potentials are often proportional to n

$f(\beta) = \lim_{n\to\infty} F_n(\beta)/n$ free energy density.

$\phi(\beta) = \lim_{n\to\infty} \cdots /n$.

$u(\beta) =$ -

$s(\beta) =$ - -

$f(\beta)$ will be always contious.

$f(\beta)$ is often analytic in some region of β

At some β_c $f(\beta)$ can be non-analytic.

Phase transition at β_c.

④ Ensemble average of an observable.

$$M(\sigma) = \sum_{i=1}^{n} \sigma_i.$$

Goal: to calculate $\lim_{n\to\infty} \langle M \rangle_\beta / n$. H

Some oracles: Given any Hamiltonian H_λ we are able to calculate $F(\beta,\lambda) = -\frac{1}{\beta} \log \int \exp\{-\beta H_\lambda(\sigma)\} \nu(d\sigma)$.

Idea: Introduce a pertubed system

$$H_\lambda(\sigma) = H(\sigma) + \lambda M(\sigma).$$

$$P_{\beta,\lambda}(d\sigma) = \frac{1}{Z(\beta,\lambda)} \exp\{-\beta H_\lambda(\sigma)\} \nu(d\sigma).$$

$$Z(\beta,\lambda) = \int \exp(\quad) \nu(d\sigma)$$

$$\Phi(\beta,\lambda) = \log Z(\beta,\lambda).$$

$$F(\beta,\lambda) = -\frac{1}{\beta} \log \Phi(\beta,\lambda).$$

$$\langle g \rangle_{\beta,\lambda} = \int_\Omega g(\sigma) P_{\beta,\lambda}(d\sigma).$$

☆ Prop:
$$\partial_\lambda F(\beta,\lambda) = \langle M \rangle_{\beta,\lambda}$$
$$\partial_\lambda F(\beta,0) = \langle M \rangle_\beta$$

Rmk: $f(\beta,\lambda) = \lim_{n\to\infty} F_n(\beta,\lambda)/n$. $\quad m(\beta,\lambda) = \lim_{n\to\infty} \langle M \rangle_{\beta,\lambda}/n$

$$\partial_\lambda f(\beta,\lambda) = m(\beta,\lambda).$$

⑤ Ensemble variance of an observable.

Lemma:
$$\partial_\lambda^2 \Phi(\beta,\lambda) = \beta^2 \left[\langle M^2 \rangle_{\beta,\lambda} - \langle M \rangle_{\beta,\lambda}^2 \right]$$
$$= \beta^2 \, Var_\beta(M)$$