Mean Field Asymptotics in Statistical Learning.

Jan 27 th.

Chapter 2
[MM 08]

Lecture 3: A. Ising models.

B. Basic concepts in statistical decision theory.

A. Ising model.

① Curie–Weiss model. Ferromagnets.

$\Omega = \{+1, -1\}^n$, $H_\lambda(\sigma) = -\frac{1}{2n} \sum_{\substack{1\le i,j\le n \\ i\ne j}} \sigma_i \sigma_j + \lambda \sum_{i=1}^n \sigma_i$

$= -\frac{1}{2n}\left(\left(\sum_{i=1}^n \sigma_i\right)^2 - n\right) + \lambda \sum_{i=1}^n \sigma_i$

$M(\sigma) = \sum_{i=1}^n \sigma_i$ $\quad P_{\beta,\lambda}(\sigma) \propto \exp\{-\beta H_\lambda(\sigma)\}$

Intersted in $m_*(\beta,\lambda) = \lim_{n\to\infty} \langle M \rangle_{\beta,\lambda} / n$

Intuition:

High temp limit: $\beta \to 0$. $T \to \infty$.

$P_{0,\lambda} = \mathrm{Unif}(\{\pm 1\})$ $\qquad m_*(0,\lambda) = 0$.

Low temp limit: $\beta \to \infty$ $\quad T \to 0$

$\lambda = 0$ $\quad \arg\min H_0(\sigma) = \{1_n, -1_n\}$ $\quad P_{\infty,0} = \frac{1}{2}\delta_{1_n} + \frac{1}{2}\delta_{-1_n}$, $m_*(\infty, 0) = 0$

$\lambda = 0+$ $\quad \arg\min H_\lambda(\sigma) = -1_n$ $\quad P_{\infty,0} = \delta_{-1_n}$ $\quad \lim_{\lambda\to 0+} m_*(\infty,\lambda) = -1$

$\lambda = 0-$ $\quad \arg\min H_\lambda(\sigma) = 1_n$ $\quad P_{\infty,0} = \delta_{1_n}$ $\quad \lim_{\lambda \to 0-} m_*(\infty,\lambda) = 1$.

Goal: $m_*(\beta,\lambda)$ $\qquad$ phase transition at β_c

a) Calculte the $f(\beta,\lambda)$

b) Calculate $\partial_\lambda f(\beta,\lambda) = m_*(\beta,\lambda)$,

c). Plot some figure.

a). $Z_n(\beta,\lambda) = \sum_{\sigma\in\{\pm1\}^n} \exp\{-\beta H_\lambda(\sigma)\}$. $\quad \phi(\beta,\lambda) = \lim_{n\to\infty} \frac{1}{n}\log Z_n(\beta,\lambda)$ $\quad f(\beta,\lambda) = -\frac{1}{\beta}\phi(\beta,\lambda)$.

$\bar m(\sigma) = \frac{1}{n}\sum_{i=1}^n \sigma_i$ $\quad$ instanteanous magnetization

$H_\lambda(\sigma) = \frac{1}{2} n (1 - \bar m(\sigma)^2) + n\lambda \bar m(\sigma)$.

$Z_n(\beta,\lambda) = \sum_{m\in \mathcal{M}_n} \sum_{\bar m(\sigma) = m} \exp\{-\beta H_\lambda(\sigma)\}$

$= \sum_{m\in\mathcal{M}_n} \mathcal{N}_n(m)\cdot \exp\{-\beta\left(\frac{1}{2}n(1-m^2) + n\lambda m\right)\}$.

$\doteq \sum_{m\in\mathcal{M}_n} \exp\{n\,\phi_{mf}(m;\beta,\lambda)\}$.

$\phi_{mf}(m;\beta,\lambda) = -\frac{\beta}{2}(1-m^2) - \beta\lambda m + \mathcal{H}\left(\frac{1+m}{2}\right)$

$\mathcal{M}_n = \{-1, \frac{-n+2}{n}, \cdots, \frac{n-2}{n}, 1\}$,

$\mathcal{N}_n(m) = |\{\sigma : \bar m(\sigma) = m\}|$

$= \binom{n}{n\frac{1+m}{2}}$

$\doteq \exp\{n\mathcal{H}(\frac{1+m}{2})\}$.

$\mathcal{H}(p) = -p\log p - (1-p)\log(1-p)$.

"$\doteq$" $\quad a_n \doteq b_n$

$\lim_{n\to\infty}\frac{1}{n}\log a_n = \lim_{n\to\infty}\frac{1}{n}\log b_n$

$\phi(\beta,\lambda) = \lim_{n\to\infty}\frac{1}{n}\log\sum_{m\in\mathcal{M}_n}\exp(n\,\phi_{mf}(m;\beta,\lambda))$

$= \max_{m\in[-1,1]} \phi_{mf}(m;\beta,\lambda)$

$f(\beta,\lambda) = -\frac{1}{\beta}\phi(\beta,\lambda) = \min_{m\in[-1,1]}\left[\frac{1}{2}(1-m^2) + \lambda m - \frac{1}{\beta}\mathcal{H}\left(\frac{1+m}{2}\right)\right]$

Convex envelop thm

Implicit differentiation

$f(\lambda) = \min_m f(m;\lambda)$

$m_* = \arg\min_m f(m;\lambda)$

$f'(\lambda) = \partial f(m;\lambda)\big|_{m=m_*}$

b) $m_*(\beta,\lambda) = \partial_\lambda f(\beta,\lambda) = m_*(\beta,\lambda)$

$m_* = \arg\min$

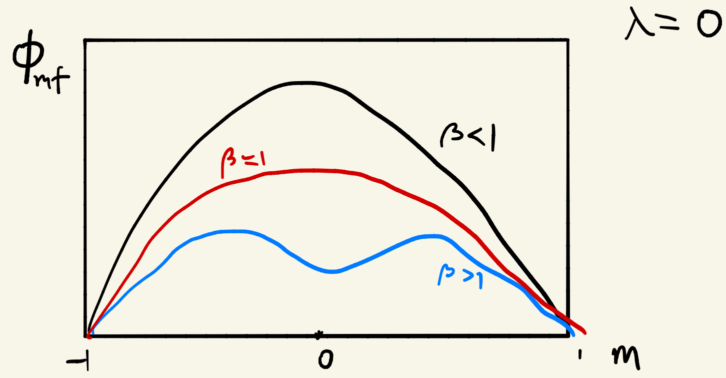
$m_* = \tanh(\beta m_* - \beta\lambda)$.

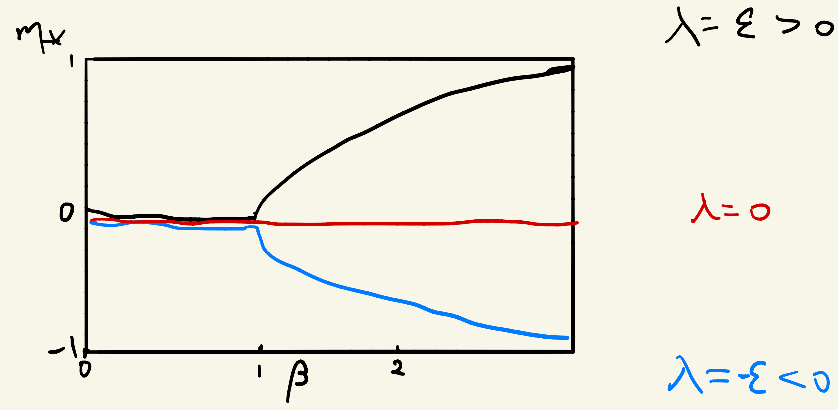
c).

Property: $\underline{\bar{m}(\sigma)}$, $\underline{\sigma \sim P_{\beta,0}}$ $\quad \mu_{\bar{m}} \approx \frac{1}{2}\delta_{m_*} + \frac{1}{2}\delta_{-m_*}$

$P_{\beta_0}(\sigma) = P_{\beta,0}(-\sigma) \qquad \langle M \rangle_{\beta,0} = 0$

$$m_*(\beta,\lambda) = \arg\max_m \phi_{mf}(m;\beta,\lambda) = \arg\max_m \left\{ -\frac{\beta}{2}(1-m^2) - \beta\lambda m + h\left(\frac{1+m}{2}\right)\right\}.$$





Phase transition at $\beta = 1$.

Property of ferromagnet.

For $\beta < 1$. (High temperature)

$\lim_{\lambda \to 0+} m_*(\beta,\lambda) = \lim_{\lambda \to 0-} m_*(\beta,\lambda) = 0$. No residual magnetization.

For $\beta > 1$ (Low temperature)

$\lim_{\lambda \to 0+} m_*(\beta,\lambda) > 0 \qquad \lim_{\lambda \to 0-} m_*(\beta,\lambda) < 0$, Residual magnetization.

② Spin-glass model. Sherrington-Kirkpatrick

$\Omega = \{\pm 1\}^n \qquad H_\lambda(\sigma) = -\sum_{i,j=1}^{n} J_{ij}\sigma_i\sigma_j - \frac{\lambda}{n}\sum_{i,j=1}^{n}\sigma_i\sigma_j$

$J_{ij} \overset{i.i.d.}{\sim} N(0,\frac{1}{n}),\ i<j \qquad J_{ii} \overset{i.i.d.}{\sim} N(0,\frac{2}{n})$.

Free entropy: $\Phi_n(\beta,\lambda) \equiv \log \sum_{\sigma\in\{\pm 1\}^n} \exp\{-\beta H_\lambda(\sigma)\}$.

$$\phi(\beta,\lambda) = \lim_{n\to\infty} \Phi_n(\beta,\lambda)/n$$

$$\phi(\beta,\lambda) = \inf_{q,b}\left[\phi_{mf}(q,b;\lambda,\beta) = \frac{1}{4}\beta^2(1-q)^2 - \frac{1}{2}\beta\lambda b^2 + \mathbb{E}_{g\sim N(0,1)}\left[\log 2\cosh(\beta\lambda b + \sqrt{q}\,g)\right]\right]$$

for some region of β and λ.

$\mathbb{Z}_2$ synchronization.

Statistical physics		Statistical decision theory.
Configuration space Ω		Parameter space Θ.
Reference measure ν_0		Bayes prior Q.
random Hamiltonian $H: \Omega \to \mathbb{R}$	$\longleftrightarrow$	log-Likelihood funtion $\log \mathbb{P}_\theta(Y)$
Gibbs-measure $\propto \exp\{-\beta H(\sigma)\}\nu_0(d\sigma)$		Bayes posterior $\propto \mathbb{P}_\sigma(Y)\, Q(d\sigma)$
Average configuration under Gibbs measure: finite β	$\longleftrightarrow$	Bayes estimator
Average configuration under Gibbs measure: $\beta = \infty$	$\longleftrightarrow$	M-estimator (MLE)
Ensemble average of an obserable (choose proper obserable): finite β	$\longleftrightarrow$	Risk of Bayes estimator
Ensemble average of an obserable (choose proper obserable): $\beta = \infty$	$\longleftrightarrow$	Risk of M-estimator.
Free energy $-\frac{1}{\beta}\log\int \exp\{-\beta H(\sigma)\}\nu_0(d\sigma)$	$\longleftrightarrow$	Mutual information $I(\text{observation}; \text{signal})$.