# Mean Field Asymptotics in Statistical Learning.

Feb 24th

## Lecture 10: Replica methods and $\mathbb{Z}_2$ synchronization.

### ① $\mathbb{Z}_2$ synchronization.

Signal: $\theta \in \mathbb{R}^n$, $\theta_i \sim_{i.i.d.} \mathrm{Unif}(\{\pm 1\})$. $\lambda \geq 0$.

Observation: $Y = \frac{\lambda}{n}\theta\theta^T + W \in \mathbb{R}^{n\times n}$. $W \sim GOE(n)$.

Estimator: $\hat{\theta}(Y) = \sum_{\sigma \in \mathbb{Z}_2^n} \sigma P_{\beta,\lambda}(\sigma)$

Gibbs measure: $P_{\beta,\lambda}(\sigma) \propto \exp\{\beta \langle \sigma, Y\sigma\rangle\}$.

$\begin{cases} \text{MLE} & \beta = \infty \\ \text{Bayes} & \beta = \lambda/2 \end{cases}$

$\psi : \mathbb{R}^2 \to \mathbb{R}$

Limiting observables:

$$m_*(\beta,\lambda) \equiv \lim_{n\to\infty} \mathbb{E}[\langle \sum_{i=1}^n \psi(\sigma_i, \theta_i)\rangle_{\beta,\lambda}]/n.$$

$$s_*(\beta,\lambda) \equiv \lim_{n\to\infty} \mathbb{E}[\sum_{i=1}^n \psi(\langle\sigma_i\rangle_{\beta,\lambda}, \theta_i)]/n.$$

Formalism: (Our goal today).

(A) $$m_*(\beta,\lambda) \equiv \lim_{n\to\infty} \mathbb{E}[\langle \sum_{i=1}^n \psi(\sigma_i,\theta_i)\rangle_{\beta,\lambda}]$$

$$= \mathbb{E}_{G,\theta}\left[\mathbb{E}_{\bar\sigma\sim D(\tanh(2\beta\lambda\mu_*\theta + 2\beta\sqrt{q_*}G))}[\psi(\bar\sigma,\theta)]\right]$$

$G \sim N(0,1)$, $\theta \sim \mathrm{Unif}(\{\pm 1\})$. $D(m)$ is a distribution.

$\mathbb{E}_{\bar\sigma\sim D(m)}[\bar\sigma] = m$ $\quad$ $\mathrm{supp}\, D(m) = \{\pm 1\}$.

Actually: $\lim_{n\to\infty} \sum_{i=1}^n \psi(\sigma_i,\theta_i)/n = \mathbb{E}[\cdots]$

Interperetation: $\frac{1}{n}\sum_{i=1}^n \delta_{(\sigma_i,\theta_i)} \longrightarrow$ Law of $(\bar\sigma,\theta)$

$\theta \sim \mathrm{Unif}(\{\pm1\})$ $\quad G\sim N(0,1)$. $\quad \bar\sigma \sim D(\tanh(2\beta\lambda\mu_*\theta + 2\beta\sqrt{q_*}G))$

(B) $$s_*(\beta,\lambda) \equiv \lim_{n\to\infty} \mathbb{E}[\sum_{i=1}^n \psi(\langle\sigma_i\rangle_{\beta,\lambda},\theta_i)]/n.$$

$$= \mathbb{E}_{G,\theta}[\psi(\tanh(2\beta\lambda\mu_*\theta + 2\beta\sqrt{q_*}G), \theta)].$$

Interpretation: $\frac{1}{n}\sum_{i=1}^n \delta_{(\langle\sigma_i\rangle_{\beta,\lambda},\theta_i)} \longrightarrow$ Law of $(\mathbb{E}[\bar\sigma],\theta)$.

Example 1: $\psi(a,\theta) = (a-\theta)^2$.

(B) $\Rightarrow$ $\lim_{n\to\infty} \mathbb{E}[\|\langle\sigma\rangle_{\beta,\lambda} - \theta\|_2^2]/n = \mathbb{E}_{G,\theta}[(\tanh(2\beta\lambda\mu_*\theta + 2\beta\sqrt{q_*}G) - \theta)^2]$

Example 2: $\psi(a,\theta) = (\mathrm{sign}(a) - \theta)^2$

(B) $\Rightarrow$ $\lim_{n\to\infty} \mathbb{E}[\|\mathrm{sign}\langle\sigma\rangle_{\beta,\lambda} - \theta\|_2^2]/n = \mathbb{E}_{G,\theta}[(\mathrm{sign}(\tanh(\cdot)) - \theta)^2]$.

$$\begin{cases} \mu_\star = \mathbb{E}[\theta \tanh(2\beta\lambda\mu_\star\theta + 2\beta\sqrt{q_\star}G)] \\ q_\star = \mathbb{E}[\tanh(2\beta\lambda\mu_\star\theta + 2\beta\sqrt{q_\star}G)^2]. \end{cases}$$

Example 3: $\psi(a,\theta) = a\theta$

$$(B) \Rightarrow \lim_{n\to\infty} \mathbb{E}\Big[\sum_{i=1}^n \langle\sigma_i\rangle_{\beta,\lambda}\theta_i/n\Big] = \mathbb{E}[\theta\tanh(2\beta\lambda\mu_\star\theta + 2\beta\sqrt{q_\star}G)].$$

$$= \lim_{n\to\infty}\mathbb{E}[\langle\langle\sigma\rangle_{\beta,\lambda},\theta\rangle]/n = \mu_\star.$$

$$\lim_{n\to\infty}\mathbb{E}[\langle\langle\sigma\rangle_{\beta,\lambda},\theta\rangle^2]/n^2 = \mu_\star^2$$

$$\left(\begin{array}{c} \langle\sigma\rangle_{\beta,\lambda} = 0. \\ P_{\beta,\lambda}(\sigma) \propto \exp\{\beta\langle\sigma, Y\sigma\rangle\} \\ \sigma_1 = 1. \end{array}\right)$$

② Free energy trick and replica trick for $m_\star(\beta,\lambda)$.

$$\Omega = \mathbb{Z}_2^n, \quad \nu_0 = \mathrm{Unif}$$

perturbation.

$$H_{\lambda,h}(\sigma) = -\langle\sigma, W\sigma\rangle - \lambda\langle\sigma,\theta\rangle^2/n - h\sum_{i=1}^n \psi(\sigma_i,\theta_i).$$

$$Z_n(\beta,\lambda,h) = \int_\Omega \exp\{-\beta H_{\lambda,h}(\sigma)\}\,\nu_0(d\sigma).$$

$$\varphi(\beta,\lambda,h) = \lim_{n\to\infty}\frac{1}{n}\mathbb{E}[\log Z_n(\beta,\lambda,h)].$$

$$m_\star(\beta,\lambda) = \frac{1}{\beta}\partial_h\varphi(\beta,\lambda,h)\Big|_{h=0}.$$

We will show: only correct in some $(\beta,\lambda)$

$$\varphi(\beta,\lambda,h) = \operatorname*{ext}_{\mu,q}\, u(q,\mu;\beta,\lambda,h)$$

$$u(q,\mu;\beta,\lambda,h) = -\beta\lambda\mu^2 + \beta^2(1-q)^2 + \mathbb{E}_{G,\theta}\Big\{\log\mathbb{E}_\sigma\Big[\exp\{2\beta\lambda\mu\sigma\theta + 2\beta\sqrt{q}G\sigma + \beta h\psi(\sigma,\theta)\}\Big]\Big\}$$

$$(G,\theta,\sigma) \sim N(0,1)\times\mathrm{Unif}(\{\pm1\})\times\mathrm{Unif}(\{\pm1\})$$

Replica trick:

a) $S(k,\beta,\lambda,h) \equiv \lim_{n\to\infty}\frac{1}{n}\log\mathbb{E}[Z_n(\beta,\lambda,h)^k]$. The $n$ limit.

b) $\varphi(\beta,\lambda,h) \equiv \lim_{k\to0}\frac{1}{k}S(k,\beta,\lambda,h)$. The $k$ limit.

c) $m_\star(\beta,\lambda) \equiv \frac{1}{\beta}\partial_h\varphi(\beta,\lambda,h)\Big|_{h=0}$ The $h$ derivative.

a) The $n$ limit :

$$\mathbb{E}[Z_n^k] = \mathbb{E}\left[\left(\int_\Omega \exp\{-\beta H_{\lambda,h}(\sigma)\}\, \nu_0(d\sigma)\right)^k\right]$$

$$= \mathbb{E}\left[\int_{\Omega^{\otimes k}} \exp\{-\beta \sum_{a=1}^k H_{\lambda,h}(\sigma^a)\} \prod_{a=1}^k \nu_0(d\sigma)\right].$$

$$= \int_{\Omega^{\otimes k}} \exp\{\beta(\sum_{a=1}^k \lambda \frac{\langle\theta,\sigma^a\rangle^2}{n} + \underline{h \sum_{a=1}^k \sum_{i=1}^n \psi(\sigma_i^a, \theta_i)})\}$$

$$\times \underbrace{\mathbb{E}[\exp\{\beta \sum_{a=1}^k \langle\sigma^a, W\sigma^a\rangle\}]}_{E} \underline{\prod_{a=1}^k \nu_0(d\sigma^a)}$$

$$E = \exp\{\beta^2 \sum_{ab=1}^k \langle\sigma^a,\sigma^b\rangle^2/n\}.$$

$$\mathbb{E}[Z_n^k]$$

$$= \int_{\Omega^k} \exp\{\beta \sum_{a=1}^k \lambda \frac{\langle\theta,\sigma^a\rangle^2}{n} + \beta^2 \sum_{ab=1}^k \langle\sigma^a,\sigma^b\rangle^2/n$$

$$\underline{+\beta h \sum_{a=1}^k \sum_{i=1}^n \psi(\sigma_i^a,\theta_i)}\} \prod \nu_0(d\sigma^a).$$

$$1 = \int \prod \delta(\langle\theta,\sigma^a\rangle - n q_{0a}) \prod \delta(\langle\sigma^a,\sigma^b\rangle - n q_{ab}) \prod dq_{0a} \prod dq_{ab}$$

$$= \exp\{\beta\lambda n \sum_{a=1}^k q_{0a}^2 + \beta^2 n \sum_{ab=1}^k q_{ab}^2\} \times \mathrm{Ent}$$

$$\mathrm{Ent} \equiv \int_{\Omega^k} \prod_{a=1}^k \delta(\langle\theta,\sigma^a\rangle - n q_{0a}) \prod \delta(\langle\sigma^a,\sigma^b\rangle - n q_{ab})$$

$$\times \exp\{\beta h \sum_{a=1}^k \sum_{i=1}^n \psi(\sigma_i^a,\theta_i)\} \prod \nu_0(d\sigma^a).$$

Exercise !

$\sigma = (\sigma^0, \sigma^1, \cdots \sigma^k) \in \mathbb{R}^{k+1}$

$\sigma^a \overset{\text{iid}}{\sim} \mathrm{Unif}(\{\pm 1\})$.

$$\frac{1}{n}\log \mathrm{Ent} = \inf_{\Lambda \in \mathbb{R}^{(k+1)\times(k+1)}} \Big\{\langle Q,\Lambda\rangle/2$$

$$+ \log \underline{\mathbb{E}_\sigma}\left[\exp\{-\sum_{ab=0}^k \lambda_{ab}\sigma^a\sigma^b/2 + \beta h \sum_{a=1}^k \psi(\sigma^a,\sigma^0)\}\right]$$

$$S(k,\beta,\lambda,h) \equiv \lim_{n\to\infty} \frac{1}{n}\log \mathbb{E} Z_n^k$$

$$= \sup_Q \inf_\Lambda U(Q,\Lambda)$$

$$U(Q,\Lambda) = \beta\lambda \sum_{a=1}^k q_{0a}^2 + \beta^2 \sum_{ab=1}^k q_{ab}^2$$

$$+ \langle Q,\Lambda\rangle/2 + \log \mathbb{E}\left[\exp\{-\sum_{ab=0}^k \lambda_{ab}\sigma^a\sigma^b + \beta h \sum_{a=1}^k \psi(\sigma^a,\sigma^0)\}\right]$$

b) The $k$ limit.

$$\phi(\beta,\lambda,h) = \lim_{k\to 0} \frac{1}{k} \sup_{Q} \inf_{\Lambda} U(Q, \Lambda).$$

$$\Pi = \begin{bmatrix} 1 & 0 \\ 0 & \bar{\Pi} \end{bmatrix} \qquad \bar{\Pi} \text{ permutation.}$$

$$U(\Pi Q \Pi^T, \Pi \Lambda \Pi^T) = U(Q, \Lambda).$$

$$Q = \begin{bmatrix} 1 & \mu & \cdots & \mu \\ \mu & \ddots & & q \\ \vdots & & \ddots & \\ \mu & q & & 1 \end{bmatrix} \qquad \Lambda = \begin{bmatrix} \lambda_0 & \lambda & \cdots & \lambda_1 \\ \lambda & \lambda_2 & & \lambda_3 \\ \vdots & & \ddots & \\ \lambda_1 & \lambda_3 & & \lambda_2 \end{bmatrix}.$$

$\sigma$

$$\partial_{q_{0a}} U(Q, \Lambda) = 0 \Rightarrow \lambda_{0a} = -2\beta\lambda q_{0a}, \quad 1 \le a \le k.$$

$$\partial_{q_{ab}} U(Q, \Lambda) = 0 \Rightarrow \lambda_{ab} = -4\beta^2 q_{ab}, \quad 1 \le a \ne b \le k.$$

$$S(k,\beta,\lambda,h) = \operatorname*{ext}_{Q} U(Q)$$

$$U(Q) = -\beta\lambda \sum_{a=1}^{k} q_{0a}^2 - \beta^2 \sum_{ab=1}^{k} q_{ab}^2 + \log \mathbb{E}\Big[\exp\Big\{2\beta\lambda \sum_{a=1}^{k} q_{0a}\sigma^a\sigma^0 + 2\beta^2 \sum_{ab=1}^{k} q_{ab}\sigma^a\sigma^b + \beta h \sum_{a=1}^{k} \psi(\sigma^a, \sigma^0)\Big\}\Big]$$

$q_{0a} = \mu, \qquad 1 \le a \le k$

$q_{ab} = q, \qquad 1 \le a \ne b \le k.$

$$U(k,\mu,q) = -\beta\lambda k\mu^2 - \beta^2(k + k(k-1)q^2) + 2\beta^2(1-q)k + \log\mathbb{E}\Big[\exp\Big\{2\beta\lambda\mu \sum_{a=1}^{k}\sigma^a\sigma^0 + \underline{2\beta^2 q \sum_{ab=1}^{k}\sigma^a\sigma^b} + \beta h \sum_{a=1}^{k}\psi(\sigma^a,\sigma^0)\Big\}\Big]$$

$$u(\mu, q; \beta,\lambda,h) = \lim_{k\to 0} \frac{1}{k} U(k,\mu,q).$$

Trick 1: $\mathbb{E}[e^{\lambda G \sum_{a=1}^{k} \sigma^a}] = \exp\{\frac{\lambda^2}{2}(\sum_{a=1}^{k}\sigma^a)^2\} = \exp\{\frac{\lambda^2}{2}\sum_{ab=1}^{k}\sigma^a\sigma^b\}.$

$$U(k,\mu,q) = -\beta\lambda k\mu^2 - \beta^2(k + (k-1)q^2) + 2\beta^2(1-q)k + \log \mathbb{E}_{\underline{G,\theta,\sigma}}\Big[\exp\Big\{2\beta\lambda\mu\sum_{a=1}^{k}\sigma^a\theta + 2\beta\sqrt{q}\,\boxed{G}\sum_{a=1}^{k}\sigma^a + \beta h\sum_{a=1}^{k}\psi(\sigma^a,\theta)\Big\}\Big].$$

$G \sim N(0,1)$. $\quad \theta \sim \mathrm{Unif}(\{\pm 1\}) \qquad \sigma = (\sigma^1, \cdots \sigma^k) \sim \mathrm{Unif}(\{\pm1\}^k).$

$$= -\beta\lambda k\mu^2 - \beta^2(k + (k-1)q^2) + 2\beta^2(1-q)k$$

$$+ \log \mathbb{E}_{G,\theta}\left[\left(\mathbb{E}_\sigma \exp\{2\beta\lambda\mu\sigma\theta + 2\beta\sqrt{q}\,G\sigma + \beta h\psi(\sigma,\theta)\}\right)^k\right]$$

$G \sim N(0,1)$ $\quad\theta \sim \mathrm{Unif}(\{\pm 1\})$ $\quad\sigma \sim \mathrm{Unif}(\{\pm 1\})$

Trick 2: $\lim_{k\to 0} \frac{1}{k} \log \mathbb{E}[Z(X)^k] = \mathbb{E}[\log Z(X)].$

$$\lim_{k\to 0} \frac{1}{k} U(k,\mu,q) \;\overset{=\,u(q,\mu;\beta,\lambda,h)}{=}\; -\beta\lambda\mu^2 + \beta^2(1-q)^2$$

$$+ \mathbb{E}_{G,\theta}\left[\log \mathbb{E}_\sigma\left[\exp\{2\beta\lambda\mu\sigma\theta + 2\beta\sqrt{q}\,G\sigma + \beta h\psi(\sigma,\theta)\}\right]\right]$$

$$Z(X) = \mathbb{E}_\sigma\left[\exp\{2\beta\lambda\mu\sigma\theta + 2\beta\sqrt{q}\,G\sigma + h\beta\psi(\sigma,\theta)\}\right].$$

$$\varphi(\beta,\lambda,h) = \operatorname*{ext}_{q,\mu}\; u(q,\mu;\beta,\lambda,h)$$

c) The $h$ derivative.

$$m_*(\beta,\lambda) = \frac{1}{\beta}\partial_h \varphi(\beta,\lambda,h) = \frac{1}{\beta}\partial_h u(q,\mu;\beta,\lambda,0)\Big|_{\substack{q=q_* \\ \mu=\mu_*}}.$$

$$\partial_h u(q,\mu;\beta,\lambda,0) = \mathbb{E}_{G,\theta}\left[\frac{\mathbb{E}_\sigma[\exp\{2\beta\lambda\mu\sigma\theta + 2\beta\sqrt{q}\,G\sigma\}\psi(\sigma,\theta)]}{\mathbb{E}_\sigma[\exp\{2\beta\lambda\mu\sigma\theta + 2\beta\sqrt{q}\,G\sigma\}]}\right].$$

$$= \mathbb{E}_{G,\theta}\left[\mathbb{E}_{\bar\sigma\sim D(\tanh(2\beta\lambda\mu + 2\beta\sqrt{q}\,G))}[\psi(\sigma,\theta)]\right]$$

$m_*(\beta,\lambda)$.