Mean Field Asymptotics in Statistical Learning.

Feb 22th.

Lecture 9 Replica methods.
Spiked GOE matrix
$\mathbb{Z}_2$ synchronization.

① The spiked GOE matrix and the free energy approach.

Let $u \in \mathbb{S}^{n-1} = \{ x \in \mathbb{R}^n : \|x\|_2 = 1 \}$.

$\lambda \in \mathbb{R}_+$, $W \sim GOE(n)$.

$$Y = \lambda u u^T + W \in \mathbb{R}^{n \times n}.$$

We are interested in calculating

a) $$\varphi(\lambda) \equiv \lim_{n\to\infty} \mathbb{E}\Big[\sup_{\sigma \in \mathbb{S}^{n-1}} \langle \sigma, Y\sigma \rangle \Big].$$

b) $$m(\lambda) \equiv \lim_{n\to\infty} \mathbb{E}\big[\langle v_{max}(Y), u \rangle^2 \big].$$

where $v_{max}(Y) = \arg\max_{\sigma \in S^{n-1}} \langle \sigma, Y \sigma \rangle$.

The free energy approach:

[Crisanti, Sommers, 1992
The spherical p-spin interaction spin glass model: the statics].
p = 2, only replica symmetric part.
No spike

$$\Omega = S^{n-1}, \quad \nu_0 = \text{Unif}$$

$$H_\lambda(\sigma) \equiv -n \langle \sigma, W\sigma \rangle - n\lambda \cdot \langle \sigma, u \rangle^2.$$

$$Z_n(\beta,\lambda) \equiv \int_{S^{n-1}} \exp\{-\beta H_\lambda(\sigma)\} \nu_0(d\sigma)$$

$$\Phi_n(\beta,\lambda) \equiv \log Z_n(\beta,\lambda).$$

$$\varphi(\beta,\lambda) \equiv \lim_{n\to\infty} \mathbb{E}[\log Z_n(\beta,\lambda)]/n \leftarrow ?$$

$$\varphi(\lambda) \equiv \lim_{\beta\to\infty} \frac{1}{\beta} \varphi(\beta,\lambda)$$

$$m(\lambda) \equiv \varphi'(\lambda).$$

[Sherrington-Kirkpatrick, 1975]

② The replica trick.

Our goal: $\varphi(\beta,\lambda) \equiv \lim_{n\to\infty} \mathbb{E}[\log Z_n(\beta,\lambda)]/n$.

Lemma: $\mathbb{E}[\log Z] = \lim_{k\to 0} \frac{1}{k} \log \mathbb{E}[Z^k]$.

$$\Rightarrow \varphi(\beta,\lambda) = \lim_{k\to 0} \lim_{n\to\infty} \frac{1}{nk} \log \mathbb{E}[Z_n(\beta,\lambda)^k].$$

4 steps

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

b) $\varphi(\beta,\lambda) = \lim_{k\to 0} \frac{1}{k} S(k,\beta,\lambda)$, The k limit

c) $\varphi(\lambda) = \lim_{\beta\to\infty} \frac{1}{\beta} \varphi(\beta,\lambda)$. The β limit

d) $m(\lambda) = \varphi'(\lambda)$ The λ differentiation.

a) The n limit.

Lemma: for $k \in \mathbb{N}_+$

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

$$= \sup_{\substack{Q \in \mathbb{R}^{(k+1)\times(k+1)} \\ \mathrm{diag}(Q)=1 \\ Q \succeq 0}} \{ U(Q) \}.$$

where $U(Q) = \beta\lambda \sum_{i=1}^{k} q_{0i}^2 + \beta^2 \sum_{ij=1}^{k} q_{ij}^2 + \frac{1}{2}\log\det(Q)$.

$$Q = (q_{ij})_{0\le i,j\le k} \in \mathbb{R}^{(k+1)\times(k+1)}.$$

Derivation:

$$\mathbb{E}[Z_n(\beta,\lambda)^k] = \mathbb{E}\left[\left(\int_{S^{n-1}} \exp\{-\beta H_\lambda(\sigma)\} \nu_0(d\sigma)\right)^k\right].$$

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

$$= \int_{(S^{n-1})^{\otimes k}} \mathbb{E}\left[\exp\{\beta n \sum_{a=1}^{k} (\lambda \langle u, \sigma^a\rangle^2 + \langle \sigma^a, W\sigma^a\rangle)\}\right] \prod_{a=1}^{k} \nu_0(d\sigma^a)$$

$$= \int_{(S^{n-1})^{\otimes k}} \exp\{\beta n \sum_{a=1}^{k} \lambda \langle u,\sigma^a\rangle^2\} \times \underbrace{\mathbb{E}\left[\exp\{\beta n \sum_{a=1}^{k} \langle \sigma^a, W\sigma^a\rangle\}\right]}_{E} \prod_{a=1}^{k} \nu_0(d\sigma^a)$$

$$W = (G+G^T)/\sqrt{2n}, \qquad G \in \mathbb{R}^{n\times n}, \qquad G_{ij} \overset{i.i.d.}{\sim} N(0,1).$$

$$E = \mathbb{E}\left[\exp\left\{\beta n \sum_{a=1}^{k} \langle \sigma^a, (G+G^T)\sigma^a\rangle/\sqrt{2n}\right\}\right]$$

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

See last lecture for full derivation

$$\mathbb{E}[Z_n(\beta,\lambda)^k]$$

$$= \int_{(S^{n-1})^k} \exp\left\{\beta n \sum_{a=1}^{k} \lambda \langle u, \sigma^a\rangle^2 + \beta^2 n \sum_{ab=1}^{k} \langle \sigma^a, \sigma^b\rangle^2\right\} \prod_{a=1}^{k} \nu_0(d\sigma^a).$$

$$\left(\text{Plug in } 1 = \int \prod_{a=1}^{k} \delta(\langle u, \sigma^a\rangle - q_{0a}) \prod_{ab=1}^{k} \delta(\langle \sigma^a, \sigma^b\rangle - q_{ab}) \prod dq_{ab}\right)$$

$$= \int \prod_{0\le a<b\le k} dq_{ab} \exp\left\{\beta\lambda n \sum_{a=1}^{k} q_{0a}^2 + \beta^2 n \sum_{ab=1}^{k} q_{ab}^2\right\}$$

$$\times \underbrace{\int_{(S^{n-1})^k} \prod_{a=1}^{k} \delta(\langle u, \sigma^a\rangle - q_{0a}) \prod_{1\le a<b\le k} \delta(\langle \sigma^a, \sigma^b\rangle - q_{ab}) \prod_{a=1}^{k} \nu_0(d\sigma^a)}_{\mathrm{Ent}}$$

Laplace method

$$\doteq \sup_{\substack{Q\succeq 0\\ q_{ii}=1}} \exp\left\{\beta\lambda n \sum_{a=1}^{k} q_{0a}^2 + \beta^2 n \sum_{ab=1}^{k} q_{ab}^2\right\} \times \mathrm{Ent}$$

$$Q = (q_{ij})_{0\le i,j\le k}.$$

$$\mathrm{Ent} = \int_{S^{n-1}} \nu_0(d\sigma^0) \int_{(S^{n-1})^k} \prod_{a=1}^{k} \delta(\langle \sigma^0, \sigma^a\rangle - q_{0a}) \prod_{ab=1}^{k} \delta(\langle \sigma^a, \sigma^b\rangle - q_{ab}) \prod_{a=1}^{k} \nu_0(d\sigma^a)$$

$$= \int_{S^{n-1}(\sqrt{n})} \nu_0(d\bar\sigma^0) \int_{(S^{n-1}(\sqrt{n}))^k} \prod_{a=1}^{k} \delta(\langle \bar\sigma^0, \bar\sigma^a\rangle - n q_{0a}) \prod_{ab=1}^{k} \delta(\langle \bar\sigma^a, \bar\sigma^b\rangle - n q_{ab}) \prod_{a=1}^{k} \nu_0(d\bar\sigma^a)$$

$$= \mathbb{P}(\bar Q(\bar\sigma) \approx Q) \doteq \exp\left\{n \tfrac{1}{2}\log\det(Q)\right\}.$$

$$\mathbb{E}\left[Z_n(\beta,\lambda)^k\right] = \sup_{\substack{Q\succeq 0\\ Q_{ii}=1}} \exp\left\{ n\left(\beta\lambda \sum_{a=1}^{k} q_{0a}^2 + \beta^2 \sum_{ab=1}^{k} q_{ab}^2 + \frac{1}{2}\log\det(Q)\right)\right\}.$$

$$S(k,\beta,\lambda) \doteq \sup_{\substack{Q\succeq 0\\ Q_{ii}=1}} U(Q)$$

$$U(Q) = \beta\lambda \sum_{a=1}^{k} q_{0a}^2 + \beta^2 \sum_{a,b=1}^{k} q_{ab}^2 + \frac{1}{2}\log\det(Q).$$

b). The k limit

$$\varphi(\beta,\lambda) = \lim_{k\to 0} \frac{1}{k} S(k,\beta,\lambda) \doteq \lim_{k\to 0} \frac{1}{k} \sup_{\substack{Q\succeq 0\\ Q_{ii}=1}} U(Q)$$

Two problems : ① The formula is derived for integer k.

② k is the dimension of Q matrix

No closed form solution for argmax.

Approach : ① Just ignore this fact

② Guess a form of the solution (ansatz).

Let $\Pi = \begin{bmatrix} 1 & 0 \\ 0 & \bar{\Pi} \end{bmatrix} \in \mathbb{R}^{(k+1)\times(k+1)}$ where $\bar{\Pi}$ is a permutation metrix.

$\Rightarrow$ $U(Q) = U(\Pi Q \Pi^T)$ $\Rightarrow$ $\exists\, Q_*$ stationery s.t. $Q_* = \Pi Q_* \Pi^T$ $\forall\, \bar{\Pi}$ permutation

A first reasonable guess : the supremum is taken at

Replica symmetric ansatz :

$$Q = \begin{bmatrix} 1 & \mu & \cdots & \mu \\ \mu & 1 & & q \\ \vdots & & \ddots & \\ \mu & q & & 1 \end{bmatrix}$$

i.e. $q_{0a} = \mu$, $1 \le a \le k$,

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

$$\Rightarrow U(Q) = \beta\lambda k\mu^2 + \beta^2 k + \beta^2 k(k-1)q^2$$
$$+ \frac{1}{2}\left[\log\left(1 + k[(q-\mu^2)/(1-q)]\right) + k\log(1-q)\right].$$

$$\log\det\begin{bmatrix} 1 & \mu 1^T \\ \mu 1 & A \end{bmatrix} = \log\det(A - \mu^2 1 1^T)$$

$$\log\det(Q) = \log\det\left((1-q) I_k + (q-\mu^2) 1 1^T\right)$$
$$= \log\det\left(I_k + [(q-\mu^2)/(1-q)]\, 1 1^T\right) + k\log(1-q)$$
$$= \log\left(1 + k[(q-\mu^2)/(1-q)]\right) + k\log(1-q)$$

$$\varphi(\beta,\lambda) = \underline{\lim_{k\to 0}}\ \underline{\sup_{q,\mu}}\ \frac{1}{k} U(k,q,\mu) \quad \text{exchange ?}$$

$$= \operatorname*{ext}_{q,\mu}\ \lim_{k\to 0} \frac{1}{k} U(k,q,\mu) = \operatorname*{ext}_{q,\mu}\ u(q,\mu;\beta,\lambda)$$

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

$$= \beta\lambda\mu^2 + \beta^2(1-q^2) + \frac{q-\mu^2}{2(1-q)} + \frac{1}{2}\log(1-q).$$

$$\Rightarrow \quad \partial_\mu u = 2\beta\lambda\mu - \frac{\mu}{1-q} \qquad \begin{pmatrix} -1<\mu<1 \\ -1<q<1 \end{pmatrix}$$

$$\partial_q u = -2\beta^2 q - \frac{\mu^2}{2(1-q)^2} + \frac{1}{2(1-q)^2}.$$

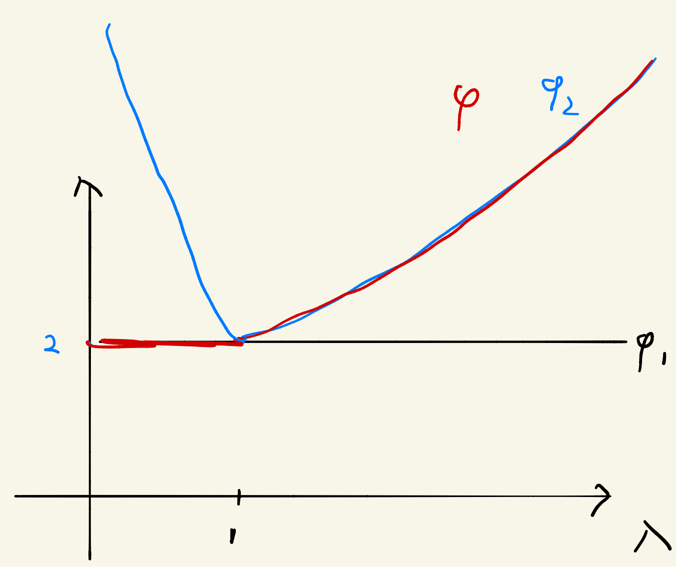
Two branch of solutions

- $\mu_1 = 0$, $q_1 = 1-\frac{1}{2\beta}$ $\Rightarrow$ $\varphi_1(\beta,\lambda) = u(\mu_1,q_1;\beta,\lambda) = 2\beta - \frac{3}{4} - \frac{1}{2}\log(2\beta).$
- $\mu_2 = \left(\left(1-\frac{1}{\lambda^2}\right)\left(1-\frac{1}{2\beta\lambda}\right)\right)^{\frac{1}{2}}$, $q_2 = 1-\frac{1}{2\lambda\beta}$ $\Rightarrow$ $\varphi_2(\beta,\lambda) = \beta\left(\lambda+\frac{1}{\lambda}\right) - \left(\frac{1}{4\lambda^2}+\frac{1}{2}\right) - \frac{1}{2}\log(2\lambda\beta).$

c) Large β limit.

$$\varphi_1(\lambda) = \lim_{\beta\to\infty} \frac{1}{\beta}\varphi_1(\beta,\lambda) = 2.$$

$$\varphi_2(\lambda) = \lim_{\beta\to\infty} \frac{1}{\beta}\varphi_2(\beta,\lambda) = \lambda + \frac{1}{\lambda}.$$



Select a branch :

Note
$$\varphi(\lambda) = \lim_{n\to\infty} \mathbb{E}\left[\sup_{\sigma\in S^{n-1}} \langle \sigma, (\lambda u u^T + W)\sigma \rangle\right]$$
$$= \lim_{n\to\infty} \mathbb{E}\left[\lambda_{\max}(\lambda u u^T + W)\right]$$

should be non-decreasing in λ.

and $\lim_{\lambda\to\infty} \varphi(\lambda) = \infty.$

- When $\lambda \le 1$, $\varphi_2 = \lambda + \frac{1}{\lambda}$ is decreasing. We need to select branch $\varphi(\lambda) = 2$.
- When $\lambda > 1$, $\varphi_1 = 2$ stay constant, We need to take $\varphi(\lambda) = \lambda + \frac{1}{\lambda}$.

This gives $\varphi(\lambda) = \begin{cases} 2 & , \quad \lambda \le 1 \\ \lambda + \frac{1}{\lambda} & , \quad \lambda > 1. \end{cases}$

known as BBP phase transition

Remark 1: How to interpret $Q_* = \underset{\substack{Q \succeq 0 \\ Q_{ii}=1}}{\arg\max}\, U(Q)$

and $(\mu_*, q_*) = \underset{q,\mu}{\arg\text{ext}}\, u(q,\mu)$?

Define $\mu_{\beta,\lambda}(d\sigma) \propto \exp\{-\beta H_\lambda(\sigma)\}$.

Then $\cancel{\neq}$ $\mathbb{E}[\langle \|\sigma\|_2^2 \rangle_{\beta,\lambda}]$.

$$\mathbb{E}\, \|\langle \sigma \rangle_{\beta,\lambda}\|_2^2 = \mathbb{E} \langle \int \sigma\, \mu_{\beta,\lambda}(d\sigma), \int \sigma\, \mu_{\beta,\lambda}(d\sigma) \rangle$$

$$= \mathbb{E} \int_{\Omega^2} \langle \sigma^1, \sigma^2 \rangle\, \mu_{\beta,\lambda}(d\sigma^1)\, \mu_{\beta,\lambda}(d\sigma^2)$$

$$= \lim_{k\to 0} \frac{2}{k(k-1)}\, \mathbb{E}\left[\int_{\Omega^r} \sum_{1\le a<b\le r} \langle \sigma^a, \sigma^b \rangle \prod_{a=1}^{k} \mu_{\beta,\lambda}(d\sigma^a) \right].$$

$$\Rightarrow \lim_{n\to\infty} \mathbb{E}\, \|\langle \sigma \rangle_{\beta,\lambda}\|_2^2 \overset{\text{Laplace method}}{=} \lim_{k\to 0} \frac{2}{k(k-1)} \sum_{1\le a<b\le k} Q^*_{ab} \overset{\text{replica trick}}{=} q_* .$$

where $Q^* = \underset{\substack{Q \succeq 0 \\ Q_{ii}=1}}{\arg\max}\, U(Q)$.

$q^* = \underset{\mu, q}{\arg\text{ext}}\, u(\mu, q)$.

Furthermore

$$\mathbb{E}[\langle \langle \sigma, u \rangle \rangle_{\beta,\lambda}] = \lim_{k\to 0} \frac{1}{k}\, \mathbb{E}\left[\int_{\Omega^r} \sum_{a=1}^{r} \langle \sigma^a, \sigma^0 \rangle \prod_{a=1}^{k} \mu_{\beta,\lambda}(d\sigma^a) \right]$$

$$\doteq \lim_{k\to 0} \frac{1}{k} \sum_{a=1}^{k} Q^*_{0a} = \mu_* .$$

So the estimation error of estimator $\langle \sigma \rangle_{\beta,\lambda}$ gives

$$\mathbb{E}\left[\|\langle \sigma \rangle_{\beta,\lambda} - u\|_2^2 \right]$$

$$= \mathbb{E}[\|\langle \sigma \rangle_{\beta,\lambda}\|_2^2] - 2\,\mathbb{E}[\langle \langle \sigma, u \rangle \rangle_{\beta,\lambda}] + 1$$

$$= 1 - 2\mu_* + q_* .$$

$\lambda > 1$.

MLE : $\beta = \infty$, $\quad \mu_* = \sqrt{1 - \frac{1}{\lambda^2}}$. $\quad q_* = 1$

$$\Rightarrow \quad \mathbb{E}[\| \hat{\theta}_{sp} - u \|_2^2] \doteq 2 - 2 \cdot \sqrt{1 - \frac{1}{\lambda^2}}.$$

Bayes estimator : $\beta = \frac{\lambda}{2}$. $\quad \mu_* = 1 - \frac{1}{\lambda^2} \quad q_* = 1 - \frac{1}{\lambda}$.

$$\Rightarrow \quad \mathbb{E}[\| \hat{\theta}_{Bayes} - u \|_2^2] \doteq \frac{2}{\lambda^2} - \frac{1}{\lambda}.$$

Remark 2: Replica symmetric ansatz doesn't always hold.

Some conditions that it holds.

① It holds whenever the overlap of two replicas $\langle \sigma^1, \sigma^2 \rangle$ concentrates,

where $(\sigma^1, \sigma^2) \sim \mathbb{E}[\mu_{\beta,\lambda}(\cdot) \times \mu_{\beta,\lambda}(\cdot)]$.

② It holds when the measure $\mu_{\beta,\lambda}$ is log-concave or when $H_\lambda(\sigma)$ is convex.

When replica symmetric ansats does not hold, there is a hierachy of ansats which are called k steps of replica symmetric breaking.

Remark 3 : Some tips in the calculation

① First step in a) is to get rid of the expectation.

② Always keep the typical order of the exponent $O(n)$.

$\sigma_i \in S^{n-1}$ $\quad \delta(\langle \sigma_i, \sigma_j \rangle - q) = \frac{1}{2\pi} \int_{\mathbb{R}} \exp\{ n i \lambda \langle \sigma_i, \sigma_j \rangle - i n \lambda q \} d\lambda$

$\sigma_i \in \{\pm 1\}^n$ $\quad \delta(\langle \sigma_i, \sigma_j \rangle / n - q) = \frac{1}{2\pi} \int_{\mathbb{R}} \exp\{ i \lambda \langle \sigma_i, \sigma_j \rangle - i n \lambda q \} d\lambda$.

③ If stuck, introduce δ functions and use delta identity formula. You will find the formula to be longer but more tractable.

Example: $L: \mathbb{R} \to \mathbb{R}$, $X \in \mathbb{R}^{n \times d}$, $X_{ij} \overset{i.i.d.}{\sim} N(0,1)$.

$$\mathbb{E}_X \left[\int_{\mathbb{R}^d} \exp\left\{ - \sum_{i=1}^n L((X\sigma/\sqrt{n})_i) \right\} \nu_0(d\sigma) \right]$$

$$= \mathbb{E}_X \left[\int_{\mathbb{R}^n} d\tau \int_{\mathbb{R}^d} \exp\left\{ - \sum_{i=1}^n L((X\sigma/\sqrt{n})_i) \right\} \delta(X\sigma/\sqrt{n} - \tau) \nu_0(d\sigma) \right]$$

$$= \mathbb{E}_X \int_{\mathbb{R}^n} d\tau \int_{\mathbb{R}^d} \exp\left\{ - \sum_{i=1}^n L(\tau_i) \right\} \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} d\lambda \exp\left\{ i \langle X\sigma/\sqrt{n}, \lambda \rangle - i \langle \tau, \lambda \rangle \right\} \nu_0(d\sigma)$$

$$= \int_{\mathbb{R}^n} d\tau \int_{\mathbb{R}^n} d\lambda \int_{\mathbb{R}^d} \nu_0(d\sigma) \exp\left\{ - \sum_{i=1}^n L(\tau_i) - i \langle \tau, \lambda \rangle \right\} \mathbb{E}\left[\exp\left\{ i \langle X\sigma/\sqrt{n}, \lambda \rangle \right\} \right]$$

Remark 4: References.

[Crisanti, Sommers, 1992] first replica calculation without spike.

[Baik, Ben Arous, Peche, 2004] Rigorous proof of spiked covariance model

[Peche, 2004] Rigorous proof of spiked GOE matrix.

analyzing density formula of spiked GOE.

HCIZ integral.

Other approach: Stieltjes transform.

AMP.

Defer this to lecture 10.

[Sherrington, Kirkpatrick, Solvable model of a spin glass, 1975].

③ Bayes estimation in $\mathbb{Z}_2$ synchronization.

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

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

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

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

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

Interested in (for some sufficiently smooth test function $\psi: \mathbb{R}^2 \to \mathbb{R}$).

$\lim_{n\to\infty} \mathbb{E}\left[\sum_{i=1}^n \psi(\hat{\theta}_i, \theta_i)\right]/n$ (Not the ensemble average of observable)

We instead define $m_*(\lambda) \equiv \lim_{n\to\infty} \mathbb{E}\left[\int \sum_{i=1}^n \psi(\sigma_i, \theta_i) p(\sigma|Y) d\sigma\right]/n$.

Free energy trick.

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

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

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

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

$m_*(\lambda) = \frac{1}{\beta}\partial_h \varphi(\beta,\lambda,h)\big|_{\beta=\lambda/2, h=0}$, Lecture 5.

② Replica method.

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) $\varphi(\beta,\lambda,h) = \lim_{k\to 0}\frac{1}{k} S(k,\beta,\lambda,h)$. The k limit.

c) $m_*(\lambda) \equiv \partial_h \varphi(\beta,\lambda,h)\big|_{\beta=\lambda/2, h=0}$ The h derivative.

a) The n limit

the same as before, except the entropy term.

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

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

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

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

$$\mathbb{E}\big[Z_n(\beta, \lambda, h)^k \big]$$

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

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

$$\Big(\text{Plug in } 1 = \int_{\Omega^k} \prod_{a=1}^{k} \delta\big(\langle u, \sigma^a \rangle - n q_{0a} \big) \prod_{ab=1}^{k} \delta\big(\langle \sigma^a, \sigma^b \rangle - n q_{ab} \big) \prod d q_{ab} \Big)$$

$$= \sup_{\substack{Q \succeq 0 \\ Q_{ii} = 1 \\ (x_a)_{a \in [k]}}} \exp\Big\{ \beta \lambda n \sum_{a=1}^{k} q_{0a}^2 + \beta^2 n \sum_{ab=1}^{k} q_{ab}^2 + h \sum_{a=1}^{k} x_a \Big\} \times \mathrm{Ent}$$

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

$$\exp\Big\{ \beta h \sum_{a=1}^{k} \sum_{i=1}^{n} \psi(\sigma_i^a, \theta_i) \Big\} \prod_{a=1}^{k} \nu_0(d\sigma^a)$$ Large deviation calculation

Exercise !

$$\frac{1}{n} \log \mathrm{Ent} = \inf_{\Lambda} \Big\{ \langle Q, \Lambda \rangle / 2 + \log \mathbb{E}\Big[\exp\Big\{ - \sum_{a,b=0}^{k} \lambda_{ab} \sigma_a \sigma_b / 2 + \beta h \sum_{a=1}^{k} \psi(\sigma_a, \sigma_0) \Big\} \Big] \Big\}$$

where $(\sigma_a)_{0 \le a \le k} \sim \mathrm{Unif}(\{\pm 1\})$

$$\Rightarrow S(k, \beta, \lambda, h) = \sup_{\substack{Q \succeq 0 \\ Q_{ii} = 1}} \inf_{\Lambda} U(Q, \Lambda)$$

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

$$+ \langle Q, \Lambda \rangle / 2 + \log \mathbb{E}\Big[\exp\Big\{ - \sum_{ab=0}^{k} \lambda_{ab} \sigma_a \sigma_b / 2 + \beta h \sum_{a=1}^{k} \psi(\sigma_a, \sigma_0) \Big\} \Big].$$

Difference : ① Additional ψ term.

② The expectation cannot be calculated explicitly.

b). The k limit

$$\varphi(\beta,\lambda,h) = \lim_{k\to 0} \frac{1}{k} \sup_{\substack{Q\succeq 0 \\ Q_{ii}=1}} \inf_{\Lambda} U(Q, X, \Lambda, \rho).$$

Replica symmetric ansatz,

$$\Pi = \begin{bmatrix} 1 & \\ & \overline{\pi} \end{bmatrix} \Rightarrow U(Q,\Lambda) = U(\Pi Q \Pi^T, \Pi \Lambda \Pi^T).$$

$$Q = \begin{bmatrix} 1 & \mu & \cdots & \mu \\ \mu & 1 & & q \\ \vdots & & \ddots & \\ \mu & q & & 1 \end{bmatrix} \qquad \begin{aligned} q_{0a} &= \mu & 1 &\le a \le k \\ q_{ab} &= q & 1 &\le a \ne b \le k. \end{aligned}$$

$$\Lambda = \begin{bmatrix} \lambda_0 & \lambda_1 & \cdots & \lambda_1 \\ \lambda_1 & \lambda_2 & & \lambda_3 \\ \vdots & & \ddots & \\ \lambda_1 & \lambda_3 & & \lambda_2 \end{bmatrix}$$

Take derivative w.r.t. q_{0a}, q_{ab}.

$$\lambda_{0a} = -2\beta\lambda q_{0a}$$

$$\lambda_{ab} = -4\beta^2 q_{ab}.$$

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

$$+ \log \mathbb{E}\left[\exp\left\{2\beta\lambda \sum_{a=1}^{k} q_{0a}\sigma_a\sigma_0 + 2\beta^2 \sum_{ab=1}^{k} q_{ab}\sigma_a\sigma_b + \beta h \sum_{a=1}^{k} \psi(\sigma_a,\sigma_0)\right\}\right]$$

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

$$+ \log \mathbb{E}_\sigma \left[\exp \left\{ \beta \lambda \mu \sum_{a=1}^{k} \sigma_a \sigma_0 + 2\beta^2 (1-q) k + 2\beta^2 q \sum_{ab=1}^{k} \sigma_a \sigma_b + \beta h \sum_{a=1}^{k} \psi(\sigma_a, \sigma_0) \right\} \right]$$

A trick! $G \sim N(0, 1)$

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

$$+ \log \mathbb{E}_{\sigma, G} \left[\exp \left\{ 2\beta \lambda \mu \sum_{a=1}^{k} \sigma_a \sigma_0 + 2\beta \sqrt{q}\, G \sum_{a=1}^{k} \sigma_a + \beta h \sum_{a=1}^{k} \psi(\sigma_a, \sigma_0) \right\} \right]$$

$\sigma, \theta \overset{i.i.d.}{\sim} \mathrm{Unif}\{\pm 1\}$

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

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

$$\left(\lim_{k \to 0} \frac{1}{k} \log \mathbb{E}_G \left[Z(G)^k \right] = \mathbb{E}_G \left[\log Z(G) \right] \right)$$

$$\Rightarrow \lim_{k \to 0} \frac{1}{k} U(\mu, q) = -\beta\lambda\mu^2 \underbrace{- \beta^2 (1 - q^2) + 2\beta^2 (1-q)}_{= \beta^2 (1-q)^2}$$

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

$$\equiv u(\mu, q; \beta, \lambda, h).$$

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

c). The h derivative.

$$\frac{1}{\beta} \partial_h \varphi(\beta, \lambda, h) \Big|_{\substack{h=0 \\ \beta = \frac{\lambda}{2}}} = \frac{1}{\beta} \partial_h u(\mu, q; \beta, \lambda, h) \Big|_{(\mu = \mu_*,\ q = q_*,\ h = 0,\ \beta = \lambda/2)}$$

$$= \mathbb{E}_{G,\theta} \left[\frac{\mathbb{E}_\sigma \left[\exp\{ \lambda^2 \mu_* \sigma\theta + \lambda^2 \sqrt{q_*}\, G \sigma \} \psi(\sigma, \theta) \right]}{\mathbb{E}_\sigma \left[\exp\{ \lambda^2 \mu_* \sigma\theta + \lambda^2 \sqrt{q_*}\, G \sigma \} \right]} \right]$$

$$(\mu_*, q_*) = \operatorname*{arg\,ext}_{\mu, q} u(\mu, q; \beta, \lambda, h) \Big|_{h=0,\ \beta = \frac{\lambda}{2}}.$$

$$u(\mu, q; \beta, \lambda, h) \Big|_{\substack{h=0 \\ \beta = \frac{\lambda}{2}}} = -\frac{\lambda^2}{2} \mu^2 + \frac{\lambda^2}{4} (1-q)^2 + \mathbb{E}_{G,\theta} \left[\log 2\cosh\left(\lambda^2 \mu \theta + \lambda \sqrt{q}\, G \right) \right]$$

$$\Rightarrow \begin{cases} \mu_* = \mathbb{E}_{G,\theta} \left[\theta \cdot \tanh\left(\lambda^2 \mu_* \theta + \lambda \sqrt{q_*}\, G \right) \right] \\ q_* = \mathbb{E}_{G,\theta} \left[\tanh\left(\lambda^2 \mu_* \theta + \lambda \sqrt{q_*}\, G \right)^2 \right] \end{cases}$$

$\psi(\sigma,\theta) = \sigma\theta$

$$\mathbb{E} \langle \hat{\theta}, \theta \rangle / n = \mathbb{E}\left[\left\langle \sum_{i=1}^{n} \sigma_i \theta_i / n \right\rangle_{\beta,\lambda,h} \right] \Big|_{\beta = \lambda/2,\ h=0}$$

$$\doteq \mathbb{E}_{G,\theta}\left[\frac{\mathbb{E}_\sigma[\exp\{\lambda^2 \mu_* \sigma\theta + \lambda^2 \sqrt{q_*} G\sigma\} \sigma\theta]}{\mathbb{E}_\sigma[\exp\{\lambda^2 \mu_* \sigma\theta + \lambda^2 \sqrt{q_*} G\sigma\}]} \right]$$

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

Remark:

We are able to derive $\lim_{n\to\infty} \mathbb{E}\left[\left\langle \sum_{i=1}^{n} \psi(\sigma_i, \theta_i)\right\rangle_{\beta,\lambda}\right]/n$ using this approach.

When $\beta = \infty$, this gives $\lim_{n\to\infty} \mathbb{E}\left[\sum_{i=1}^{n} \psi(\hat{\theta}_{ML,i}, \theta_i)\right]/n$ where $\hat{\theta}_{ML}$ is the MLE

However, we cannot derive $\lim_{n\to\infty} \mathbb{E}\left[\sum_{i=1}^{n} \psi(\hat{\theta}_{Bayes,i}, \theta_i)\right]/n$

by taking $\beta = \lambda/2$, since

$$\mathbb{E}\left[\left\langle \sum_{i=1}^{n} \psi(\sigma_i, \theta_i) \right\rangle_{\beta,\lambda}\right]/n$$

$$\overset{\times}{=} \mathbb{E}\left[\sum_{i=1}^{n} \psi(\langle \sigma_i \rangle_{\beta,\lambda}, \theta_i)\right]/n .$$