STSCI 6940:

Lecture 5 — Date: 02/23

1. $f_n(\lambda) = \frac{F_n(\lambda)}{n} = \frac{1}{n}\mathbb{E}\log\int e^{H_n(x)}\,dP_0^n(x)$.

$$f_n(\lambda) \le \mathbb{E}_{x_0}\Big[\max_{m\in T_n}\varphi_n(m, x_0)\Big] + O\left(\frac{\log n}{\sqrt{n}}\right).$$

2. $\varphi_n(m, x_0) \le \hat{F}_n(\lambda q, \lambda m, x_0) + O\left(\frac{1}{n}\right)$.

$$\hat{F}_n(\lambda q, \lambda m, x_0) = \frac{1}{n}\sum_{i=1}^{n}\psi(\lambda q, \lambda m x_{0i}) + \frac{\lambda q^2}{4} - \frac{\lambda m^2}{2}.$$

Step 3: swap $\mathbb{E}_{x_0}$ and $\max_{m\in T_n}$ $\qquad T_n = \left\{\frac{k}{n} : -n \le k \le n\right\}$

Lemma:

$$\sup_{m,q} \mathbb{P}_{x_0}\left(\left|\hat{F}_n(\lambda q, \lambda m, x_0) - \mathbb{E}\hat{F}\right| \ge t\right) \le 2e^{-C\frac{nt^2}{2}} \qquad C = C(\lambda).$$

$(r, s) \longrightarrow \psi(r, S)$.

- $\left|\frac{\partial}{\partial r}\psi\right| \vee \left|\frac{\partial}{\partial s}\psi\right| \le K$.
- $\psi(r, s) = \mathbb{E}_{z\sim N(0,1)} \log\int e^{\sqrt{r}z\varepsilon + s\varepsilon - \frac{r}{2}\varepsilon^2}\,dP_0(\varepsilon)$
- $\psi(0, 0) = 0$.

$\Rightarrow \quad |\psi(r, s)| \le K\cdot(r + |s|)$.

$\Rightarrow$ Chernoff: $\mathbb{P}_{x_0}\left(|\hat{F} - \mathbb{E}\hat{F}| \geq t\right) \leq e^{-\frac{nt^2}{\kappa(r+|s|)^2}}$

$r = \lambda q$, $S = \lambda m \cdot x_{0i}$

$m \in [-1, 1]$, $q \in (0, K)$, $\quad \frac{1}{n}\sum \psi(\lambda q, \lambda m x_{0i})$.

- $\mathbb{E} \max_{m \in T_n} (\hat{F} - \mathbb{E}\hat{F})$.

$$= \frac{1}{\gamma}\mathbb{E}\log e^{\gamma(\hat{F} - \mathbb{E}\hat{F})}.$$

$$\leq \frac{1}{\gamma}\log \mathbb{E}\max_{m \in T_n} e^{\gamma(\hat{F} - \mathbb{E}\hat{F})}.$$

$$\leq \frac{1}{\gamma}\log \sum_{m \in T_n} \underbrace{\mathbb{E}\, e^{\gamma(\hat{F} - \mathbb{E}\hat{F})}}_{\leq e^{c\frac{\gamma^2}{n}}}.$$

$$\leq \frac{\log(2n+1)}{\gamma} + \frac{c\gamma}{n} \;=\; O\left(\frac{\log n}{\sqrt{n}}\right),$$

$(\gamma = \sqrt{n})$

$$\Rightarrow \mathbb{E}\max_{m \in T_n}\hat{F} \leq \max_{m \in T_n}\mathbb{E}\hat{F} + O\left(\frac{\log n}{\sqrt{n}}\right).$$

$$f_n(x) \leq \max_{m \in T_n} \mathbb{E}\hat{F}(\lambda q, \lambda m, x_0) \; o_n(1).$$

$$\bar{F}(\lambda q, \lambda m) = \mathbb{E}\hat{F} = \mathbb{E}_{z, \varepsilon_0}\log\int e^{\sqrt{\lambda q} z \varepsilon + \lambda m \varepsilon \varepsilon_0 - \frac{\lambda q}{2}\varepsilon^2} dP_0(\varepsilon)$$

$$- \frac{\lambda m^2}{2} + \frac{\lambda q^2}{4}.$$

$\bar{q} = \bar{q}(m)$ be one minimizer of $\bar{F}(\lambda q, \lambda m)$.

upper bound:

$$\boxed{\limsup_{n \to +\infty} f_n(\lambda) \leq \underbrace{\max_{m \in [-1,+1]}\inf_{q \in (0,K)}\bar{F}(\lambda q, \lambda m)}_{\bar{\Phi}(\lambda)}.}$$

$-K \quad P_0 \quad K$ $\qquad q \le \mathrm{supp}(P_0) \qquad \phi_{RS}(\lambda)$ $\quad ||$

Lower bound: $\liminf_{n \to +\infty} f_n(\lambda) \geq \sup_{q \geq 0} \left\{ \psi(\lambda q) - \frac{\lambda q^2}{4} \right\}$

$$\psi(r) = \mathbb{E} \log \int e^{\sqrt{r} z \varepsilon + r \varepsilon \varepsilon_0 - \frac{r}{2} \varepsilon^2} dP_0(\varepsilon).$$

Final step: the upper and lower bounds match.

$$\phi_{RS}(\lambda) = \bar{\Phi}(\lambda)$$

$$\phi_{RS}(\lambda) \leq \bar{\Phi}(\lambda).$$

$$\inf_{q \in (0,K)} \bar{F}(\lambda q, \lambda m) \leq \bar{F}(\lambda |m|, \lambda m).$$

$$= \bar{\Psi}(\lambda |m|, \lambda m) - \frac{\lambda m^2}{2} + \frac{\lambda |m|^2}{4}$$

$$= \bar{\Psi}(\lambda |m|, \lambda m) - \frac{\lambda |m|^2}{4}.$$

$$\bar{\Psi}(r,s) = \mathbb{E} \log \int e^{\sqrt{r} z \varepsilon + s \varepsilon \varepsilon_0 - \frac{r}{2} \varepsilon^2} dP_0(\varepsilon).$$

$\forall s \geq 0, r \geq 0$;

Lemma: $\bar{\Psi}(r,-s) \leq \Psi(r,s)$ [E, Krzakala 2018]

$$\sup_m \inf_{q \in (0,K)} \bar{F}(\lambda q, \lambda m) \leq \sup_m \bar{\Psi}(\lambda |m|, \lambda |m|) - \frac{\lambda |m|^2}{4}.$$

$$= \phi_{RS}(\lambda). \quad \blacksquare$$

Conclusion: $\forall \lambda \geq 0$

$$\boxed{\lim_{n \to +\infty} f_n(\lambda) = \phi_{RS}(\lambda) = \sup_{q \geq 0} \left\{ \psi(\lambda q) - \frac{\lambda q^2}{4} \right\}.}$$

Recall concentration of $X_n(m)$ w.r.t. w:

- $$X_n(m) = \frac{1}{n} \log \int \mathbb{1}\{x \cdot x_0 = m\} \, e^{H_n(x)} \, dP_0^n(x).$$

$$\mathbb{E}_w \max_{m \in T_n} X_n(m) \leq \max_{m \in T_n} \mathbb{E}_w X_n(m) + o_n(1).$$

$$\mathbb{E}\left[e^{\gamma (X_n(m) - \mathbb{E} X_n(m))}\right] \leq \frac{r}{\sqrt{n}} e^{\frac{K\gamma^2}{n}}.$$

Theorem [Borell, Ibragimov, Tsirelson, Sudakov]:

Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ Lipschitz with constant L:

$$|f(x) - f(y)| \leq L \|x - y\|_2.$$

then $$\mathbb{P}\left(|f(g) - \mathbb{E}[f(g)]| \geq t \right) \leq e^{-\frac{t^2}{2L^2}}.$$

$$g \sim N(0, I_n)$$

Application to $X_n(m)$,

$$w \to X_n(m) \qquad H_n(w) = \sum_{i<j} \sqrt{\frac{\lambda}{n}} w_{ij} x_i x_j + \text{———}$$

$$\frac{\partial}{\partial w_{ij}} X_n = \frac{1}{n} \sqrt{\frac{\lambda}{n}} \cdot \frac{\int \mathbb{1}\{x \cdot x_0 = m\} \, x_i x_j \, e^{H_n(x)} \, dP_0^n(x)}{\int \mathbb{1}\{x \cdot x_0 = m\} \, e^{H(x)} \, dP_0^n(x)}.$$

$$= \frac{1}{n} \sqrt{\frac{\lambda}{n}} \langle x_i x_j \rangle.$$

$$\|\nabla_w X_n\|_2^2 = \sum_{i<j} \frac{1}{n^2} \frac{\lambda}{n} \cdot \underbrace{\langle x_i x_j \rangle^2}_{\leq 1}$$

$$= \frac{n(n-1)}{2} \cdot \frac{1}{n^2} \cdot \frac{\lambda}{n} \leq \frac{\lambda}{2n} = L^2.$$

$$\Rightarrow \mathbb{P}\left(|X_n(m) - \mathbb{E} X_n(m)| \geq t \right) \leq e^{-\frac{nt^2}{2\lambda}}.$$

$$\Rightarrow \mathbb{E}\left[e^{\gamma(X_n - \mathbb{E} X_n)} \right] = \int_0^{+\infty} \mathbb{P}\left(e^{\gamma(X_n - \mathbb{E} X_n)} \geq u \right) du$$

Consequence of the RS formula for the free energy:

$$f_n(\lambda) \xrightarrow[n\to+\infty]{} \phi_{RS}(\lambda) \quad \forall \lambda \geq 0.$$

$$\frac{1}{n} I(X_0; Y) = \frac{\lambda}{2n^2} \sum_{i<j} \mathbb{E}[X_i^2 X_j^2] - f_n(\lambda).$$

$$= \frac{\lambda}{4} \frac{n(n-1)}{n^2} - f_n(\lambda).$$

$$\frac{1}{n} I(X_0; Y) \longrightarrow \frac{\lambda}{4} - \phi_{RS}(\lambda).$$

(the replica formula for the mutual information)

- I - MMSE relation:

$$f_n'(\lambda) = \frac{1}{2} \frac{1}{n^2} \sum_{i<j} \mathbb{E}\left[\mathbb{E}[X_i X_j \mid Y]^2 \right].$$

$$\frac{d}{d\lambda} \frac{1}{n} I(X_0; Y) = \frac{1}{2n^2} \, MMSE_{P_0}(\lambda).$$

Issue: we proved convergence of f_n.
Not of f_n'.

Does f_n' converge?

observations:
1. $\lambda \to f_n(\lambda)$ is differentiable.
2. $\lambda \to f_n(\lambda)$ is convex ($f_n'' \geq 0$).
3. $f_n(\lambda) \xrightarrow[n\to+\infty]{} \phi_{RS}(\lambda) \quad \forall \lambda \geq 0$

$\Rightarrow \quad \lambda \to \phi_{RS}(\lambda)$ is convex, is differentiable on a set $D = \mathbb{R} \setminus C$, C is countable.

and $\forall \lambda \in D$, $f_n'(\lambda) \longrightarrow \phi_{RS}'(\lambda)$.

$$\phi_{RS}(\lambda) = \sup_{q \geq 0} \left\{ \psi(\lambda q) - \frac{\lambda q^2}{4} \right\} \quad (**)$$

change of var $\bar{q} = \lambda q$

$$= \sup_{\bar{q} \geq 0} \left\{ \psi(\bar{q}) - \frac{\bar{q}^2}{4\lambda} \right\} . \quad (*)$$

Envelope theorem:

$$\phi_{RS}'(\lambda) = + \frac{\bar{q}_*^2}{4\lambda^2}$$ where $\bar{q}_*$ is any maximizer of $(*)$

$$= + \frac{q_*^2}{4}$$

q_* is any maximizer of $(**)$.

$\Rightarrow$ maximizer is unique.

It follows that

- $\lim_{n \to +\infty} \frac{2}{n(n-1)} \sum_{i<j} \mathbb{E}\left[\mathbb{E}[x_i x_j | Y]^2 \right] \longrightarrow q_*^2(\lambda)$
- $\lim_{n \to +\infty} MMSE_n(\lambda) \longrightarrow 1 - q_*^2(\lambda)$.

Properties of $\lambda \to q_*(\lambda)$.

1. $\lambda \in D \to q_*(\lambda)$ is non-decreasing (ϕ_{RS} is convex)

2. $\lim_{\substack{\lambda \to 0 \\ \lambda \in D}} q_*(\lambda) = \left(\mathbb{E}_{P_0}[X_0] \right)^2 = 0$.

3. $\lim_{\substack{\lambda \to +\infty \\ \lambda \in D}} q_*(\lambda) = \mathbb{E}_{P_0}[X_0^2] = 1$.

4. $y = \sqrt{\lambda q}\, x_0 + z \qquad x_0 \sim P_0,\ z \sim N(0,1).$

$\mathrm{mmse}(\lambda q)$: MMSE of the above channel.

$$q^*(\lambda) = \mathbb{E}_{P_0}[x_0^2] - \mathrm{mmse}(\lambda q_*(\lambda)).$$

$$= 1 - \mathrm{mmse}(\lambda q_*(\lambda)).$$

Remark : $D = \{\lambda \geq 0 : \phi_{RS} \text{ is differentiable at } \lambda\}$

$= \mathbb{R} \setminus C$, C is countably many points

ϕ_{RS} is in general not everywhere differentiable.

the points of non-diff. of ϕ_{RS} : points of phase transitions.

Examples :

$P_0 = N(0,1)$.

$$\psi(\lambda q) = \mathbb{E}_{z, x_0} \log \int e^{\sqrt{\lambda q}\, z x + \lambda q\, x x_0 - \frac{\lambda q}{2} x^2} dP_0(x)$$

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

$$\phi_{RS}(\lambda) = \sup_q \{ \underline{\qquad\qquad\qquad} - \frac{\lambda q^2}{4} \}.$$

$$q_*(\lambda) = 1 - \mathrm{mmse}(\lambda q_*(\lambda)).$$

$$\mathrm{mmse}(\lambda q_*(\lambda)) = \mathbb{E}\left[\left(x_0 - \frac{\sqrt{\lambda q}}{1+\lambda q} y\right)^2\right]$$

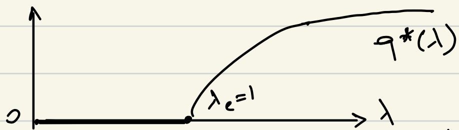
$$= 1 - 2 \frac{\sqrt{\lambda q}}{1+\lambda q} \mathbb{E}[x_0 y] + \frac{\lambda q}{(1+\lambda q)^2} \mathbb{E}[y^2]$$

$$= 1 - 2\frac{\lambda q}{1+\lambda q} + \frac{\lambda q}{1+\lambda q} \qquad \left(\text{since } \mathbb{E}[y^2] = 1+\lambda q\right)$$

$$= 1 - \frac{\lambda q}{1+\lambda q}$$

$$\boxed{q_*(\lambda) = \frac{\lambda q_*(\lambda)}{1+\lambda q_*(\lambda)}}$$

Plot:



- The reconstruction threshold:

$$\lambda_c \doteq \sup\{\lambda \geq 0 : q_*(\lambda) = 0\}.$$

(in this case $\lambda_c = 1$)

$\lambda > \lambda_c$; reconstruction is possible.

$\lambda < \lambda_c$; ———— impossible.