STSCI 6940:

Lecture 6 Date: 02/25

Examples:

1. $P_0 \doteq N(0,1)$.

$$\psi(r) = \mathbb{E} \log \int e^{\sqrt{r} y x + r x x_0 - \frac{r}{2} x^2} dP_0(x)$$

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

$$\phi_{RS}(\lambda) = \sup_{q \geq 0} \Big\{ \underbrace{\psi(\lambda q) - \frac{\lambda q^2}{4}}_{F(q)} \Big\}.$$

$$q_*(\lambda) = \frac{\lambda q_*(\lambda)}{1+\lambda q_*(\lambda)}$$ (maximize, can also obtained by direct inspection)

$q_* = 0$ is always a solution.

$$(q_* \neq 0) \Rightarrow 1 = \frac{\lambda}{1+\lambda q_*(\lambda)} \Rightarrow q_*(\lambda) = 1 - \frac{1}{\lambda}.$$

$F(0) = 0$, $F(1-\frac{1}{\lambda}) > 0$?

$$F\left(1-\frac{1}{\lambda}\right) = \frac{\lambda(1-\frac{1}{\lambda})}{2} - \frac{1}{2}\log\left(1+\lambda\left(1-\frac{1}{\lambda}\right)\right) - \frac{\lambda}{4}\left(1-\frac{1}{\lambda}\right)^2$$

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

$$= \frac{1}{2}(\lambda-1)\left(1 - \frac{\lambda-1}{2\lambda}\right) - \frac{1}{2}\log\lambda$$

$$= \frac{\lambda^2 - 1}{4\lambda} - \frac{1}{2}\log\lambda \qquad \overset{(?)}{>} 0$$

$$f(\lambda) = \lambda^2 - 1 - 2\lambda\log\lambda .$$

$$f(1) = 0 .$$

$$f'(\lambda) = 2\lambda - 2\log\lambda - 2$$

$$f'(1) = 0$$

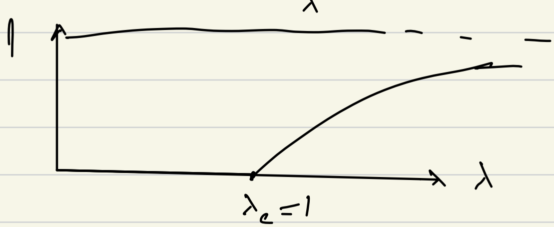
$$f''(\lambda) = 2 - \frac{2}{\lambda} . \quad \begin{cases} \geq 0 & \text{if } \lambda \geq 1 \\ < 0 & \text{o/w} . \end{cases}$$

$$\Rightarrow f'(\lambda) \begin{cases} \geq 0 & \text{if } \lambda \geq 1 \\ < 0 & \text{o/w} . \end{cases}$$

$$f(\lambda) \begin{cases} \geq 0 & \text{if } \lambda \geq 1 \\ < 0 & \text{o.w} \end{cases}$$

$$F\left(1 - \frac{1}{\lambda}\right) \begin{cases} \geq 0 & \text{if } \lambda \geq 1 \\ < 0 & \text{o/w} \end{cases}$$

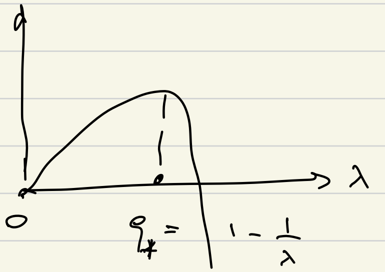
$$\eta_*(\lambda) = \begin{cases} 0 & \text{if } \lambda < 1 \\ 1 - \frac{1}{\lambda} & \text{o/w} \end{cases} .$$



$\lambda \leq \lambda_c = 1$

$q \longrightarrow F(q)$

$\lambda > \lambda_c = 1$



$0 \qquad q_* = 1 - \frac{1}{\lambda}$

Example 2: Rademacher $P_0 = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_{+1}$,

$$\psi(r) = \mathbb{E} \log\left(\frac{1}{2} e^{\sqrt{r} y - \frac{r}{2}} + \frac{1}{2} e^{-\sqrt{r} y - \frac{r}{2}}\right)$$

$$= \mathbb{E} \log \cosh(\sqrt{r} y) - \frac{r}{2}.$$

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

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

$$\mathrm{mmse}(r) = \underbrace{\mathbb{E}[x^2]}_{1} - \mathbb{E}\left[\mathbb{E}[x|y]^2\right].$$

$$\mathbb{E}[x|y] = \frac{(+1)\, e^{\sqrt{\lambda q} y - \frac{\lambda q}{2}} - 1\, e^{-\sqrt{\lambda q} y - \frac{\lambda q}{2}}}{e^{\text{—}} + e^{\text{—}}}$$

$$= \tanh(\sqrt{\lambda q} \cdot y)$$

$$q_*(\lambda) = \mathbb{E}\left[\tanh\left(\sqrt{\lambda q_*(\lambda)}\, z + \lambda q_*(\lambda) x_0\right)^2\right]$$

$$= \mathbb{E}\left[x_0 \tanh\left(\sqrt{\lambda q_*(\lambda)}\, z + \lambda q_*(\lambda) x_0\right)\right]$$

$\mathbb{E}[x_0 | y]$

$x_0 \in \{\pm 1\}$, $\Rightarrow$

$$= \mathbb{E}\left[\tanh\left(\sqrt{\lambda q_*(\lambda)}\, z x_0 + \lambda q_*(\lambda)\right)\right]$$

$$= \mathbb{E}\left[\tanh\left(\sqrt{\lambda q_*(\lambda)}\, z + \lambda q_*(\lambda)\right)\right]$$

on the other hand :

$$\mathbb{E}\left[\tanh\left(\sqrt{\lambda q_*(\lambda)}\, z + \lambda q_*(\lambda) x_0\right)^2\right].$$

$$= \mathbb{E}\left[\tanh\left(\sqrt{\lambda q_*(\lambda)}\, z + \lambda q_*(\lambda)\right)^2\right].$$

$$\Rightarrow \mathbb{E}\left[\tanh\left(\sqrt{r}\, z + r\right)^2\right] = \mathbb{E}\left[\tanh\left(\sqrt{r}\, z + r\right)\right] \quad \forall r \geq 0$$

q_* is a fixed point of the function

$$f(q) = \mathbb{E}\left[\tanh\left(\sqrt{\lambda q}\, z + \lambda q\right)\right]. \quad z \sim N(0,1)$$

$f(0) = 0$, $f(+\infty) = 1$

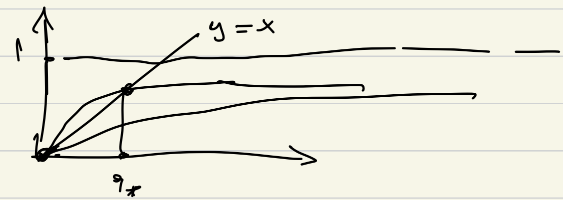
$$f'(q) = \mathbb{E}\left[\left(\frac{\sqrt{\lambda}}{2\sqrt{q}} z + \lambda\right)\left(1 - \tanh^2\left(\sqrt{\lambda q}\, z + \lambda q\right)\right)\right].$$

$$= \lambda - \lambda \mathbb{E}\left[\tanh(\ )^2\right] + \mathbb{E}\left[\frac{\sqrt{\lambda}}{2\sqrt{q}} z \, \mathrm{th}^2(\ )\right]$$

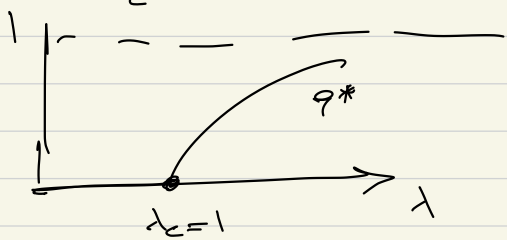
$$= \underline{\qquad\qquad\qquad\qquad} + \lambda \mathbb{E}[\, th' \, th \,].$$

$$= \lambda - \lambda \mathbb{E}[\, th^2(\,) \,] + \lambda \mathbb{E}[(1 - th^2(\,))\, th(\,)].$$

$f'(0) = \lambda$



$\lambda \gtrless 1$. $\quad \lambda_c = 1$



- Example 3: Sparse Radamacher (Three - delta prior).

$\rho \in (0,1)$

$$P_0 = (1-\rho)\,\delta_0 + \frac{\rho}{2}\,\delta_{-\frac{1}{\sqrt{\rho}}} + \frac{\rho}{2}\,\delta_{+\frac{1}{\sqrt{\rho}}}.$$

$$\psi(r) = \mathbb{E} \log\left(1 - \rho + \frac{\rho}{2}\left(e^{\sqrt{r}y} + e^{-\sqrt{r}y}\right)e^{-\frac{r}{2}}\right)$$

$$= \mathbb{E} \log\left(1 - \rho + \rho \cosh(\sqrt{r}z + r)\, e^{-r/2}\right).$$

$$q_*(\lambda) = 1 - \mathrm{mmse}(\lambda q_*(\lambda)) = \mathbb{E}[\mathbb{E}[x|y]^2]$$

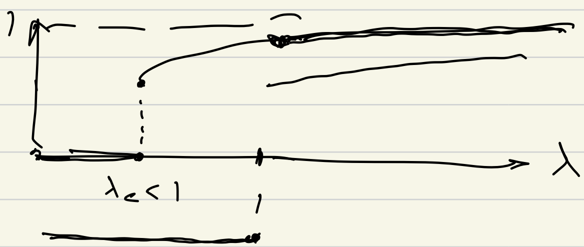
$$\mathbb{E}[x|y] = \frac{\rho \sinh(\sqrt{\lambda q}\, y)\, e^{-\frac{\lambda q}{2}}}{1 - \rho + \rho\cosh(\sqrt{\lambda q}\, y)\, e^{-\frac{\lambda q}{2}}}.$$

$\rho = 1$ $\qquad = \tanh(\sqrt{\lambda q}\, y).$

for ρ "large" : you get the same qualitative picture as $\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_{+1}$.

for ρ "small" ($\rho \leq 0.05$) :

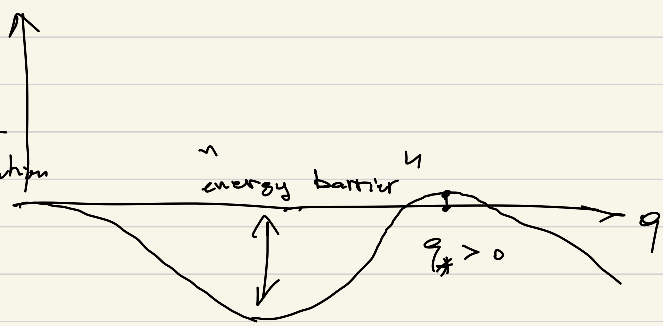
$\lambda_c < 1$



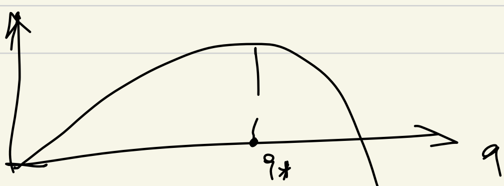
$$F(q) = \psi(\lambda q) - \frac{\lambda q^2}{4}.$$

$\lambda < \lambda_c$:
estimation is impossible

$\lambda \in (\lambda_c, 1)$
known
no polynomial time
algorithm for estimation



$\lambda \geq 1$:
Estimation is easy



- Pinning : general technique - relies on side information.

Goal : prove concentration of overlap.

- Any observation model : Draw $x_i \sim P_0$ iid $1 \leq i \leq n$

$$y \sim Q(\cdot \mid x).$$

Q is a probability kernel.

P_0 being discrete with support $[-K, K]$.

Augment observation model :

$$\forall\, 1 \leq i \leq n \text{ indep} \quad \xi_i = \begin{cases} x_i & \text{w.p } \varepsilon \\ * & \text{w.p } 1-\varepsilon \end{cases}$$

Let $L_i \in \{0,1\}$ $L_i \sim \mathrm{Ber}(\varepsilon)$:

$$L_i = \begin{cases} 1 & \text{if } x_i \text{ is revealed} \\ 0 & \text{o/w} \end{cases}$$

$$\mathbb{P}(x \mid y, \xi) \propto \prod_{i \mid L_i = 1} \mathbb{1}\{x_i = \xi_i\} \; Q(y, x) \prod_{i \mid L_i = 0} P_0(x_i)$$

$$f_{n,\varepsilon} = \frac{1}{n} \mathbb{E} \log Z_{n,\varepsilon} = \frac{1}{n} \mathbb{E} \log \Big(\sum_{x \in S^n} \qquad \Big).$$

$\varepsilon = 0$: you recover the original model
no side information.

Lemma : $\quad | f_{n,\varepsilon} - f_n | \leq \varepsilon H(P_0)$.

$$\mathbb{P}(z \mid y, L) = \sum_{x \in S^n} \mathbb{P}(z \mid L, x)\, \mathbb{P}(x \mid y).$$

$$= \sum_{x \in S^n} \prod_{i \mid L_i = 1} \mathbb{1}\{z_i = x_i\} \frac{Q(y,x) \prod_{i=1}^{n} P_0(x_i)}{Z_n}$$

$$= \frac{1}{Z_n} \left(\sum_{x \in S^n} \prod_{i \mid L_i = 1} \mathbb{1}\{z_i = x_i\}\, Q(y,x) \prod_{i \mid L_i = 0} P_0(x_i) \right) \cdot \prod_{i \mid L_i = 1} P_0(z_i)$$

$$= \frac{Z_{n,\varepsilon}}{Z_n} \underbrace{\prod_{i \mid L_i = 1} P_0(z_i)}_{\mathbb{P}(z \mid L)}.$$

$$\Rightarrow \quad f_{n,\varepsilon} - f_n = \frac{1}{n} \mathbb{E} \log \mathbb{P}(z \mid y, L) - \frac{1}{n} \mathbb{E} \log \mathbb{P}(z \mid L)$$

$$= -\frac{1}{n} H(z \mid y, L) + \frac{1}{n} H(z \mid L) \geq 0$$

$$H(z, L) \geq H(z \mid y, L).$$

$$\leq \frac{1}{n} H(z \mid L) = \frac{1}{\cancel{n}} \cdot \cancel{n} \cdot \varepsilon \cdot H(P_0)$$

$$\Rightarrow \quad \boxed{0 \leq f_{n,\varepsilon} - f_n \leq \varepsilon H(P_0).}$$

$$H(x) = H(P_x) = -\sum_{\sigma\in S} P_x(\sigma) \log P_x(\sigma) \geq 0$$

The point is $\mathbb{P}(x \mid y, \xi)^{\otimes 2}$ "simple".

satisifies concentration of overlap. ($R_{1,2}$

"decay of correlation".

Theorem: $\forall \varepsilon_0 > 0$ $\qquad \langle \cdot \rangle_{n,\varepsilon}$

$$\int_0^{\varepsilon_0} \mathbb{E} \langle (\underset{R_{1,0}}{R_{1,2}} - \underset{R_{1,0}}{\langle R_{1,2} \rangle_{n,\varepsilon}})^2 \rangle_{n,\varepsilon} \, d\varepsilon \leq c \sqrt{\frac{\varepsilon_0}{n}} .$$

$$R_{12} = x_1 \cdot x_2 = \frac{1}{n} \sum_{i=1}^{n} x_i^1 x_i^2 ,$$

$x^1, x^2 \sim \mathbb{P}(\cdot \mid y, \xi)$ indep. $\qquad y \sim Q(\cdot \mid x)$