STSCI 6940:

Lecture 7

Date: 03/02

Theorem 1: $\forall \varepsilon_0 > 0$ $\langle \cdot \rangle_{n,\varepsilon}$

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

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

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

Theorem 2: (mutual information): $\forall \varepsilon_0 > 0$

$$\int_0^{\varepsilon_0} \frac{1}{n} \sum_{i \neq j} I(x_i, x_j \mid y, \xi^{(\varepsilon)}) \, d\varepsilon \leq 2 H(P_0).$$

Thm 2 $\Rightarrow$ thm 1:

- $$R_{12} = x^1 \cdot x^2 = \frac{1}{n} \sum_{i=1}^{n} x_i^1 x_i^2 .$$

$$\langle (R_{1,2} - \langle R_{1,2} \rangle)^2 \rangle = \langle R_{1,2}^2 \rangle - \langle R_{1,2} \rangle^2$$

$$= \frac{1}{n^2} \sum_{i,j=1}^{n} \langle x_i^1 x_i^2 x_j^1 x_j^2 \rangle - \langle x_i^1 x_i^2 \rangle^2$$

$$= \frac{1}{n^2} \sum_{i,j=1}^{n} \langle x_i x_j \rangle^2 - \langle x_i \rangle^2 \langle x_j \rangle^2$$

$$\leq \frac{K^2}{n^2} \sum_{i,j=1}^{n} \left| \langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle \right|$$

$$= \frac{K^2}{n^2} \sum_{i,j=1}^{n} \Big| \sum_{\sigma,\sigma'} \mathbb{P}(x_i = \sigma, x_j = \sigma' \mid y, \mathfrak{z}^{(\varepsilon)})\, \sigma\sigma' - \mathbb{P}(x_i = \sigma \mid y, \mathfrak{z}^{(\varepsilon)})\, \mathbb{P}(x_j = \sigma' \mid y, \mathfrak{z}^{(\varepsilon)}) \cdot \sigma\sigma' \Big|$$

$$\leq \frac{K^4}{n^2} \sum_{i,j=1}^{n} 2 \cdot d_{TV}\Big(\mathbb{P}(x_i = \cdot, x_j = \cdot \mid y, \mathfrak{z}^{(\varepsilon)}),\ \mathbb{P}(x_i = \cdot \mid y, \mathfrak{z}^{(\varepsilon)}) \cdot \mathbb{P}(x_j = \cdot \mid y, \mathfrak{z}^{(\varepsilon)}) \Big)$$

Pinsker's inequality

$$\leq 2 \frac{K^4}{n^2} \sum_{i,j=1}^{n} \sqrt{\frac{1}{2} d_{KL}\big(\mathbb{P}(x_i, x_j \mid y, \mathfrak{z}^{(\varepsilon)}) \,\|\, \mathbb{P}(x_i = \cdot \mid --)\, \mathbb{P}(x_j = \cdot \mid --) \big)}$$

$$\int_0^{\varepsilon_0} \mathbb{E} \langle (R_{1,2} - \langle R_{1,2} \rangle)^2 \rangle \, d\varepsilon$$

$$\underset{\text{(Jensen)}}{\leq} 2K^4 \int_0^{\varepsilon_0} \sqrt{\mathbb{E} \frac{1}{n^2} \sum_{i,j=1}^{n} d_{KL}\big(\mathbb{P}(x_i, x_j) \,\|\, \mathbb{P}(x_i) \cdot \mathbb{P}(x_j) \big)} \, d\varepsilon$$

$$= 2K^4 \int_0^{\varepsilon_0} \sqrt{\frac{1}{n^2} \sum_{i,j=1}^{n} I(x_i, x_j \mid y, \mathfrak{z}^{(\varepsilon)})} \, d\varepsilon$$

$$\leq 2K^4 \varepsilon_0 \sqrt{\underbrace{\frac{1}{\varepsilon_0} \int_0^{\varepsilon_0} \frac{1}{n^2} \sum_{i,j=1}^{n} I(x_i, x_j \mid y, \mathfrak{z}^{(\varepsilon)}) \, d\varepsilon}_{\leq \frac{1}{n\varepsilon_0} \cdot 2 H(P_0)}}$$

$$\leq 2K^4 \sqrt{\frac{2\varepsilon_0 H(P_0)}{n}}$$

- Proof of theorem 2: (Magic lemma).

$$\forall \varepsilon_0 > 0 : \int_0^{\varepsilon_0} \frac{1}{n} \sum_{i,j} I(x_i, x_j \mid y, z^{(\varepsilon)})\, d\varepsilon \leq 2H(P_0).$$

- Reminder: $H(P_0) = -\sum_x P_0(x) \log P_0(x)$.

$$H(x,y) = H(x) + H(x \mid y).$$

$$I(x,y) = H(x) - H(x \mid y)$$

$$= H(x,y) - H(x \mid y) - H(y \mid x).$$

$$H(x \mid y) = \mathbb{E}_y H(P_{x \mid y}).$$

$$= \mathbb{E}_y \left[- \sum_\sigma P(x=\sigma \mid y) \log P(x=\sigma \mid y) \right]$$

consider:

- $\underline{\varepsilon} = (\varepsilon_1, \longrightarrow \varepsilon_n) \in [0,1]^n$

$$z_i^{(\underline{\varepsilon})} = \begin{cases} x_i & \text{w.p } \varepsilon_i \\ * & \text{w.p } 1-\varepsilon_i \end{cases} \qquad \forall\, 1 \leq i \leq n \text{ indep.}$$

$$\frac{d}{d\varepsilon_i} H(x \mid y, z^{(\underline{\varepsilon})}) \,?,$$

$$\frac{d}{d\varepsilon_i^2} H(x \mid y, z^{(\varepsilon)}),\ \frac{d^2}{d\varepsilon_i d\varepsilon_j} H(x \mid y, z^{(\varepsilon)}) \,?$$

. Claim: $\begin{cases} \dfrac{d}{d\varepsilon} H(x|y, z(\varepsilon)) = -\sum_{i=1}^{n} H(x_i | y, z^{(i)}(\varepsilon)) \\ \dfrac{d^2}{d\varepsilon^2} H(x|y, z(\varepsilon)) = \sum_{i \neq j=1}^{n} I(x_i, x_j | y, z^{(ij)}(\varepsilon)) \end{cases}$

$z^{(i)}(\varepsilon) \in \mathbb{R}^n$

$$\left(z^{(i)}(\varepsilon)\right)_j = z_j(\varepsilon) \text{ if } j \neq i$$
$$= * \quad \text{if } j = i$$

$$z^{(i)}(\varepsilon) = \begin{array}{c} \\ i \to \end{array} \begin{pmatrix} \\ * \\ \\ \end{pmatrix}$$

$$z^{(ij)}(\varepsilon) = \begin{array}{c} i \to \\ j \to \end{array} \begin{pmatrix} * \\ * \end{pmatrix}$$

Assume claim.

$$\int_0^{\varepsilon_0} \underbrace{\sum_{i,j=1}^{n} I(x_i, x_j | y, z^{(ij)}(\varepsilon))}_{= I(x_i, x_j | y, z(\varepsilon))} d\varepsilon$$

$$= \frac{d}{d\varepsilon} H(x|y, z(\varepsilon)) \Big|_{\varepsilon=\varepsilon_0} - \frac{d}{d\varepsilon} H(x|y, z(\varepsilon)) \Big|_{\varepsilon=0}$$

$$= \sum_{i=1}^{n} H(x_i | y, z^{(i)}(0)) - \underbrace{H(x_i | y, z^{(i)}(\varepsilon_0))}_{\geq 0}$$

$$\leq \sum_{i=1}^{n} H(x_i \mid y) \leq \sum_{i=1}^{n} H(x_i) = n\, H(P_0).$$ ∎

Proof of claim:

$$\frac{d}{d\varepsilon_i} H(x \mid y, z(\underline{\varepsilon})) \; ?$$

$$H(x \mid y, z(\underline{\varepsilon})) = H(x_i \mid y, z(\underline{\varepsilon})) + \underbrace{H(x \mid y, z^{(i)}(\underline{\varepsilon}), x_i)}_{\text{indep of } \varepsilon_i}$$

$$H(x_i \mid y, z(\underline{\varepsilon})) = \varepsilon_i H(x_i \mid y, z^{(i)}(\underline{\varepsilon}), x_i) + (1-\varepsilon_i) H(x_i \mid y, z^{(i)}(\underline{\varepsilon})).$$

$$H(x_i \mid y, z^{(i)}, x_i) = 0.$$

$H(x_i \mid y, z^{(i)}(\underline{\varepsilon}))$ is indep of ε_i.

$$\frac{d}{d\varepsilon_i} H(x \mid y, z(\underline{\varepsilon})) = -H(x_i \mid y, z^{(i)}(\underline{\varepsilon})).$$

[if all ε_i are equal then

$$\frac{d}{d\varepsilon} H(x \mid y, z(\varepsilon)) = -\sum_{i=1}^{n} H(x_i \mid y, z^{(i)}(\varepsilon)).$$]

$f(x_1, \ldots, x_n)$. $\quad f : \mathbb{R}^n \to \mathbb{R}$.

$$\frac{d}{dt} f(t, \ldots, t) = \sum_{i=1}^{n} \partial_i f(t, \ldots, t).$$

Mixed derivatives :

$$\frac{d^2}{d\varepsilon_i^2} H(x \mid y, z(\underline{\varepsilon})) = 0$$

For $i \neq j$:

$$\frac{d^2}{d\varepsilon_i d\varepsilon_j} H(x \mid y, z(\underline{\varepsilon})) :$$

chain rule :

$$H(x \mid y, z(\underline{\varepsilon})) = H(x_i, x_j \mid y, z(\underline{\varepsilon}))$$

$$+ \underbrace{H(x \mid y, z^{(ij)}(\underline{\varepsilon}), x_i, x_j)}_{\text{indep of } \varepsilon_i \text{ and } \varepsilon_j} .$$

$$H(x_i, x_j \mid y, z(\underline{\varepsilon})) = \varepsilon_i \varepsilon_j \underbrace{H(x_i, x_j \mid y, z^{(ij)}(\underline{\varepsilon}), x_i, x_j)}_{=0}$$

$$+ \varepsilon_i (1-\varepsilon_j) H(x_i, x_j \mid y, z^{(ij)}(\underline{\varepsilon}), x_i)$$

$$+ \varepsilon_j (1-\varepsilon_i) H(x_i, x_j \mid y, z^{(ij)}(\underline{\varepsilon}), x_j)$$

$$+ (1-\varepsilon_i)(1-\varepsilon_j) H(x_i, x_j \mid y, z^{(ij)}(\underline{\varepsilon})) .$$

$$\frac{d}{d\varepsilon_i d\varepsilon_j} H(x_i, x_j \mid y, z(\underline{\varepsilon}))$$

$$= - H(\quad x_j \mid y, z^{(ij)}(\underline{\varepsilon}), x_i)$$

$$- H(x_i \quad \mid y, z^{(ij)}(\underline{\varepsilon}), x_j)$$

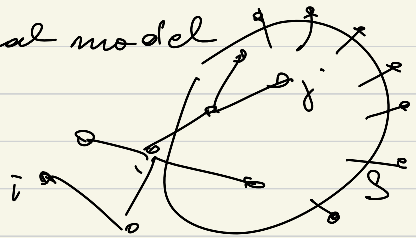
$$+ H(x_i, x_j \mid y, z^{(ij)}(\underline{\varepsilon}))$$

$$= I\left(x_i, x_j \mid y, z^{(ij)}(\varepsilon)\right).$$

[If all ε_i are equal :

$$\frac{d^2}{d\varepsilon^2} H(x \mid y, z(\varepsilon)) = \sum_{i \neq j} I(x_i, x_j \mid y, z^{(ij)}(\varepsilon)).$$]

small remark Graphical model



$$\int_0^{\varepsilon_0} \sum_{v \in S} I(\sigma_i, \sigma_{\partial S} \mid z^{(\varepsilon)})\, d\varepsilon \leq C |\partial S|.$$

specialize to K_n ; complete graph. $S = \{1, \dots, n\}$.

- Another way of revealing information:

$$y_i' = \sqrt{r}\, x_i + z_i \quad \text{iid} \quad z_i \sim N(0,1).$$
$$1 \leq i \leq n.$$

r plays the role of ε.

Free energy as a function of r.

$$\frac{d}{dr} F_n(r) = \frac{1}{2} \mathbb{E} \sum_{i=1}^{n} \mathbb{E}[x_i \mid y, y']^2$$

$$= \frac{1}{2} \sum_{i=1}^{n} \mathbb{E} \langle x_i \rangle^2 \qquad \langle x_i \rangle^2 = \langle x_i^1 x_i^2 \rangle$$

$$= \frac{n}{2} \mathbb{E} \langle R_{1,2} \rangle.$$

$$\frac{d^2}{dr^2} F_n(r) = \frac{1}{2} \mathbb{E}\left[\| \mathrm{cov}(x \mid y, y') \|_F^2 \right]$$

$$= \frac{1}{2} \mathbb{E} \sum_{i,j=1}^{n} \left(\langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle \right)^2$$

Assume $\frac{F_n}{n} \longrightarrow f_*(r)$.

$\frac{d^2}{dr^2} \frac{F_n}{n}(r) \longrightarrow f_*''(r)$ for almost every $r \geq 0$.

$$\Rightarrow \frac{1}{2n^2} \mathbb{E} \sum_{i,j=1}^{n} \left(\langle x_i x_j \rangle - \langle x_i \rangle \langle x_j \rangle \right)^2 \sim \frac{f_*''(r)}{n} \to 0$$

$$\left| \langle x_i x_j \rangle - \langle x_i \rangle \cdot \langle x_j \rangle \right|$$

$$= \Big| \sum_{\sigma, \sigma'} \mathbb{P}(x_i = \sigma, x_j = \sigma' \mid -) \sigma \sigma'$$

$$- \mathbb{P}(x_i = \sigma \mid -) \sigma \cdot \mathbb{P}(x_j = \sigma' \mid -) \sigma' \Big|$$

$$\leq k^2 \, d_{TV} \left(\mathbb{P}(x_i, x_j \mid -), \mathbb{P}(x_i = \cdot \mid -) \mathbb{P}(x_j = \cdot \mid -) \right).$$