ST06 LECTURE 4

Note Title 2/16/2006

Today

- Asymptotic Equipartition Property
- Typical Set
- Data Compression example

First : review of last lecture

Notions
- $H(X)$; $I(X;Y)$; $I(X;Y|Z)$
- $D(p \| q) = \sum_x p(x) \log \frac{p(x)}{q(x)}$
- Markov Chain : $X \to Y \to Z$

Results :
- $H(X|Y) \leq H(X)$
- $I(X;Y|Z) \leq I(X;Y)$?

No! Let $X \in_U \{0,1\}$

$Y \in_U \{0,1\}$

$Z = X + Y \pmod 2$.

then $I(X;Y) = 0$

$I(X;Y \mid Z) = 1$

- $D(p \| q) \geq 0$
- $I(X;Y) = D\left(P_{X,Y} \| P_X \cdot P_Y\right) \geq 0$
- $H(X \mid Y) = H(X) - I(X;Y) \leq H(X)$
- $H(X_1 \dots X_n) \leq \sum H(X_i)$
- $X \to Y \to Z \Rightarrow I(X;Z) \leq I(X;Y)$
- Fano's Inequality

  $X \to Y \to \tilde{X}$; $P_e = \Pr[X \neq \tilde{X}]$

$$P_e \geq \frac{H(X|Y) - 1}{\log|\Omega_x|}$$

Some Intuition / Example

(X, Y) dist as follows

w.p. $1-p$ $\quad X = Y \in_u \{0,1\}^n$

p $\quad X \in_u \{0,1\}^m$; $Y = \text{"}\lambda\text{"}$

$$\boxed{m \gg n}$$

$$H(X|Y) = (1-p) \cdot 0$$
$$+ \; p \cdot m$$

$P_e = p$

Fano's Inequality $\quad P_e \geq \frac{p \cdot m - 1}{m} \approx p.$

Today

"Convergence to typical element"

- Suppose $X_1, \ldots, X_n$ i.i.d. with dist X.

What should

$$P(x_1, \ldots x_n) \quad \text{be ?}$$

$$P(x_1 \ldots x_n) = P(x_1) \cdot P(x2) \ldots P(x_n)$$

$$\log P(x_1 \ldots x_n) = \sum \log P(x_i) \rightarrow \underbrace{E[\log P(x_i)]}_{\shortparallel \atop H(x)}.$$

Can prove this formally using law of large numbers

Weak version:

- $X_1 \dots X_n$ i.i.d $\sim p(x)$ then

$$-\frac{1}{n} \log p(X_1 \dots X_n) \longrightarrow H(X) \text{ in probability}$$

- Formally, $\forall \epsilon, \delta > 0$, $\exists n_0 < \infty$ s.t. $\forall n \geq n_0$

$$\Pr_{X_1 \dots X_n} \left[H(X) - \epsilon \leq -\frac{1}{n} \log P(X_1 \dots X_n) \leq H(X) + \epsilon \right] \geq 1 - \delta .$$

(Actually $\delta \to 0$ as $\exp(-\epsilon^2 n)$.)

Typical Set

AEP says for $(1-\delta)$ fraction of probability, el'ts satisfy

$$2^{-(H(x)+\epsilon)n} \le P(x_1 \dots x_n) \le 2^{-(H(x)-\epsilon)n}$$

Motivates defn

For $\epsilon > 0$, $n < \infty$

$$A_\epsilon^{(n)} = \left\{ (x_1 \dots x_n) \;\middle|\; 2^{- - -} \le P(x_1 \dots x_n) \le 2^{\dots} \right\}$$

Typical Set Theorem

For $n \geq n_0(\epsilon, \delta)$

1. $\Pr\left[A_\epsilon^{(n)}\right] \geq 1 - \delta$

2. $\left|A_\epsilon^{(n)}\right| \leq 2^{n(H(x)+\epsilon)}$

3. $\left|A_\epsilon^{(n)}\right| \geq 2^{n(H(x)-\epsilon)} \cdot (1-\delta)$

Example: $Z = 0$ w.p. $\frac{9}{10}$

$= 1$ w.p. $\frac{1}{20}$

$= -1$ w.p. $\frac{1}{20}$

$(Z_1 \ldots Z_n)$ contains $\frac{9(1 \pm \epsilon')}{10} \cdot n$ 0

$\frac{1}{20}(1 \pm \epsilon')n$ 1's

$\frac{1}{20}(1 \pm \epsilon')n$ "-1"'s

$A_\epsilon^{(n)} \subseteq (Z_1 \ldots Z_n) \ldots$

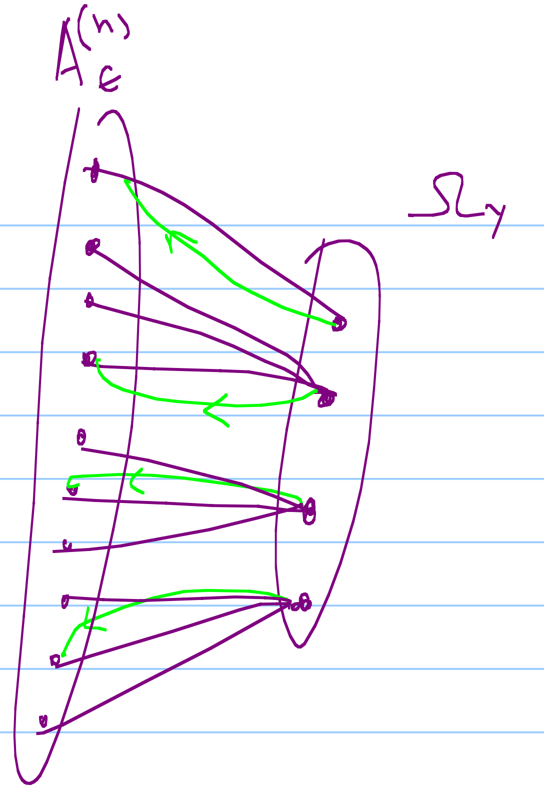
$$A_\epsilon^{(n)} \approx 2^{(H(Z) \pm \epsilon)n}$$

Application first & then proof of Typical Set Theorem.

Data Compression:

Compress $(X_1 \ldots X_n) \longrightarrow Y \in \Omega_Y$

$$P_e \geq 1 - \frac{|\Omega_Y|}{|A_\epsilon^{(n)}|}$$



Decode function maps $\Omega_y \to A_\epsilon^{(n)}$.

$$\Pr\left[\text{decoding correctly} \mid A_\epsilon^{(n)}\right]$$

$$\leq \Pr\left[\text{pts in } A_\epsilon^{(n)} \text{ that are in image of Decode}\right]$$

$$\leq |\Omega_y| \cdot \max_{(x_1 \ldots x_n) \in A_\epsilon^{(n)}} p(x_1 \ldots x_n)$$

$$\leq |\Omega_y| \cdot 2^{-(H(X)-\epsilon)n}$$

$$\leq |\Omega_y| \cdot \frac{2^{2\epsilon n}}{|A_n^{(\epsilon)}|} .$$

$$\Pr[\text{decoding correctly}]$$

$$\leq \delta + \frac{|\Omega_Y|}{|A_n^{(\epsilon)}|} \cdot 2^{\epsilon n}$$

$$\xrightarrow{\quad\quad\quad} 0$$

Proof of typical set theorem

① Same as AEP.

② $\Pr\left[A_\epsilon^{(n)}\right] \leq 1$

$$\Rightarrow \left|A_\epsilon^{(n)}\right| \cdot \left\{ \min_{\bar{x} \in A_\epsilon^{(n)}} p(\bar{x}) \right\} \leq 1$$

$$\overset{\text{\tiny ||}}{\smile}$$

$$2^{-(H(X)+\epsilon)n}$$

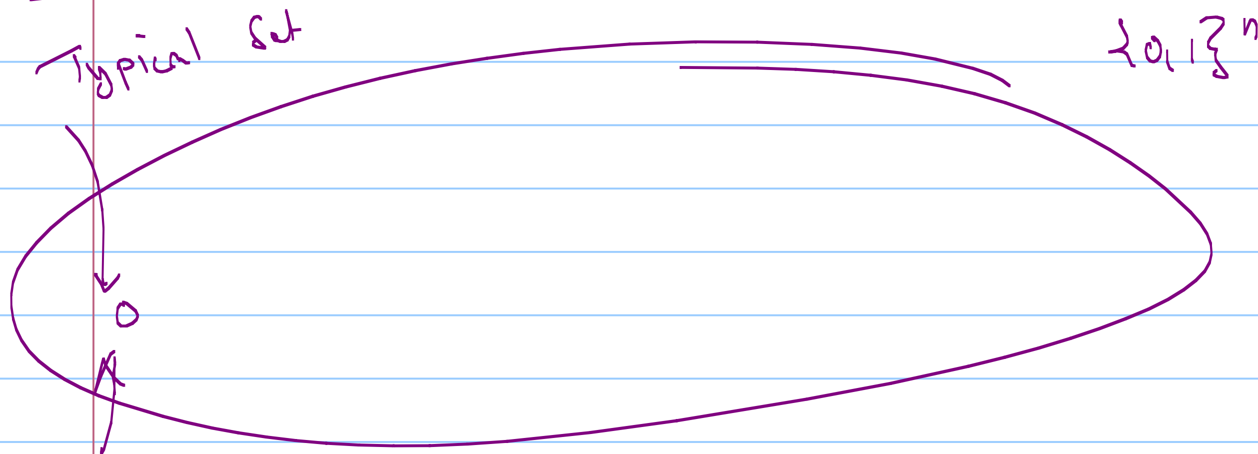
$$\Rightarrow \quad \left| A_\epsilon^{(n)} \right| \leq 2^{(H(X)+\epsilon)\cdot n}$$

③ $$\Pr\left[A_\epsilon^{(n)}\right] \geq 1-\delta$$

$$\Rightarrow \quad \left|A_\epsilon^{(n)}\right| \cdot \max_{x \in A} p(\bar{x}) \geq 1-\delta$$

$$\Rightarrow \quad \left|A_\epsilon^{(n)}\right| \geq (1-\delta)\cdot 2^{(H(X)-\epsilon)\cdot n}$$

Pictures of Typical Set



All probability is concentrated here

if B is any set s.t.

$$\Pr[B] \geq 1-\delta$$

$$\Rightarrow \quad |B \cap A_\epsilon^{(n)}| \geq 2^{(H(x)-\epsilon')n} \cdot (1-2\delta)$$

Proof

- $\Pr[B \cap A_\epsilon^{(n)}] \geq 1-2\delta$

- But
$$\Pr[B \cap A_\epsilon^{(n)}] \leq |B \cap A_\epsilon^{(n)}| \cdot 2^{-(H(x)-\epsilon)n}$$

QED.