INFORMATION THEORY IN CS

CS229r LECTURE 1

1/29/2019

This Course

1 BASIC INFORMATION THEORY

- NOTIONS : ENTROPY, INFORMATION, DIVERGENCE
- TOOLS : CHAIN RULES, INEQUALITIES,

2 - 6 : APPLICATIONS

E.g.
- Compression of Information
- Design of codes + algorithms.
- Analyze Communication Complexity
- Analyze Limiting behavior of games.
- Limits of Data Structures
- Tight analysis in differential privacy

⋮

TODAY : INFORMATION THEORY IN COMBINATORICS.

Basics (informally):

X - random variable

P_X - distribution of X.

- Suppose Alice + Bob know P_X; Alice also knows X
- How many bits (in expectation) does Alice have to send to Bob to convey X?

"Entropy of X" → $H(X)$ roughly measures this.

- Conditional Entropy: $(X,Y) \sim P_{XY}$
  - Alice + Bob know X & P_{XY}
  - How many bits (in expectation over X,Y) does Alice send to Bob to convey Y?

  "Conditional Entropy of $Y|X$" denoted $H(Y|X)$

- Key relationships if $X \in \mathcal{X}$

① $H(X) \leq \log|\mathcal{X}|$

② $H(X,Y) = H(X) + H(Y|X)$

③ $H(Y|X) \leq H(Y)$.

TODAY'S APPLICATION

SHEARER'S LEMMA (BABY VERSION?):

Notation
$[N] = \{1, 2, .. N\}$

Lets • $F \subseteq [N]^d$

• for $S \subseteq [d]$ let $F_S \subseteq [N]^{|S|}$

$S = \{i_1, .. i_s\}$

$= \{(x_{i_1}, x_{i_2}, .. x_{i_s}) \mid (x_1 .. x_d) \in F\}$

$=$ "Projection of F to S".

• ~~Theore~~ Lemma: $[d=3]$ version

$$|F|^2 \leq |F_{\{1,2\}}| \cdot |F_{\{1,3\}}| \cdot |F_{\{2,3\}}|$$

• (d, k)-version ("exercise")

$$|F|^{\binom{d-1}{k-1}} \leq \prod_{\substack{S \subseteq [d] \\ |S| = k}} |F_S| .$$

Motivation

1 $d=3$, $k=1$ version

Lemma asserts $|F| \le |F_1| \cdot |F_2| \cdot |F_3|$

But this is obvious since

$$F \subseteq F_1 \times F_2 \times F_3$$

& so $|F| \le |F_1| \cdot |F_2| \cdot |F_3|$

$k=2$ quite non-trivial!

2 Lemma is tight:

Suppose $F = F_1 \times F_2 \times \cdots F_d$

then $F_S = \prod_{j \in S} F_j$

$$\prod_{\substack{S \subseteq [d] \\ |S| = k}} \prod_{j \in S} F_j = \prod_{j=1}^{n} F_j^{\binom{d-1}{k-1}} = F^{\binom{d-1}{k-1}}$$

Our Proof [due to "Radhakrishnan"]

Idea: find distribution (random variable) (X,Y,Z) supported on F

s.t. • $H(X,Y,Z) = \log |F|$

• Achive $H(X,Y) \leq \log |F_{12}|$

$H(Y,Z) \leq \log |F_{23}|$

$H(X,Z) \leq \log |F_{13}|$

• Suffices to Prove:

$$2 \log |F| = 2H(X,Y,Z) \leq H(X,Y) + H(X,Z) + H(Y,Z).$$

$$\leq \log F_{12} + \log F_{23} + \log F_{13}$$

$$H(X,Y) = H(X) + H(Y|X)$$

$$H(X,Z) = H(X) + H(Z|X)$$

$$H(Y,Z) = H(Y) + H(Z|Y)$$

$$2H(X) \qquad \geq 2H(Y|X) \qquad \geq 2H(Z|X,Y)$$

MORAL:

- Somewhere out there, in the axioms there was a deep theorem.

- Which One?

Rest of Course

- Fill in basics (relatively short)

- Get to applications

Your tasks

1. Scribe work - Sign up! (15%)
2. Problem Sets - 40%
3. Project - 30%
4. Participation - 15%

- Remember to sign up on Piazza.
- Follow Class website

http://madhu.cs.harvard.edu/courses/Spring2019