round elimination & triangular discrimination

amir yehudayoff

outline

round elimination [...Miltersen-Nisan-Safra-Wigerson...]

triangular discrimination [Topsøe]

round elimination

“model”

computational process with R rounds

input X

in round i some data $M_i(X, M_{<i})$ is recorded

M_R should provide some info on X

problem: lower bounds on R...

round elimination

construct $X^{(R)}$ and $f^{(R)}$ so that solution for them in R rounds yields solution for $X^{(R-1)}$ and $f^{(R-1)}$ in $R-1$ round

$X^{(0)}$ and $f^{(0)}$ are non trivial

error increases in every step

example: communication complexity

alice gets X and bob gets Y

they talk: $M_1(X), M_2(Y, M_1), M_3(X, M_1, M_2), \ldots, M_r$

round elimination

level 0: alice gets X, bob gets Y, compute $f(X, Y)$

level 1: alice gets $(X_1, \ldots, X_n)$, bob gets $X_{<j}, Y$, compute $f(X_j, Y)$

...

suggestion: do not eliminate rounds

bound amount of info collected

for all i

$$\mathop{\mathbb{E}}_{M_{\leq i}} dist\left(p_{X^{(R-i)}} \middle| m_{\leq i}, p_{X^{(R-i)}}\right) \leq i \cdot \epsilon$$

choice of *dist* is important

triangular discrimination

measures of "distance" between distributions p, q are useful

each measure due to unique properties

f**-divergence [Csiszar, Morimoto, Ali, Silvey]**

$$D_f(p||q) = \sum_{\omega} q(\omega) f\left(\frac{p(\omega)}{q(\omega)}\right)$$

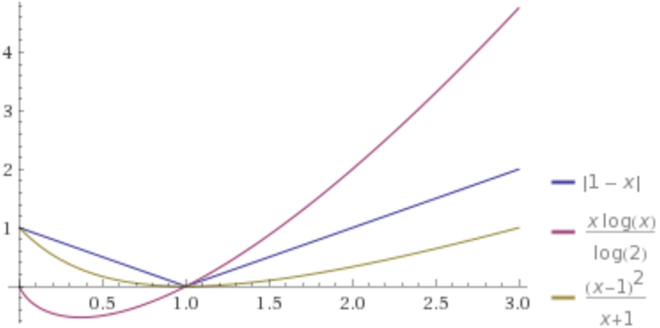
with f convex so that $f(1) = 0$

examples

ℓ_1 distance $|p - q|_1$ with $|1 - x|$

KL-divergence $D(p||q)$ with $x \log_2 x$

triangular discrimination $\Delta(p, q)$ with $\frac{(1-x)^2}{1+x}$



properties

non-negativity $D_f(p||q) \geq 0$

convexity $D_f(p||q)$ is convex in (p, q)

data processing $D_f(p_X||p_Y) \geq D_f(p_{g(X)}||p_{g(Y)})$

relations

[Pinsker] $|p - q|_1 \leq \sqrt{2D(p||q)}$

extremely useful in information theoretic proofs

simple $\Delta(p,q) \leq |p - q|_1 \leq \sqrt{2\Delta(p,q)}$

$\Delta(p,q) = \sum_\omega \frac{(p(\omega)-q(\omega))^2}{p(\omega)+q(\omega)}$

[Topsøe] $\Delta(p,q) \leq 2D(p||q)$

dual/operational meaning

ℓ_1 & statistical distance

$$|p - q|_1 = \max_{\|g\|_\infty \le 1} \left| \mathop{\mathbb{E}}_{p} g - \mathop{\mathbb{E}}_{q} g \right|$$

Δ & ℓ_2

$$\Delta(p, q) = \max_{\mathbb{E}_p g^2 + \mathbb{E}_q g^2 \le 1} \left(\mathop{\mathbb{E}}_{p} g - \mathop{\mathbb{E}}_{q} g \right)^2$$

applications

Δ was recently used

$\star$ construct group homomorphisms
[Erschler, Karlsson]

$\star$ study harmonic functions on groups
[Benjamini, Duminil-Copin, Kozma, Yadin]

$\star$ Gromov's theorem on groups of polynomial growth
[Ozawa]

an example

assume X takes values in $\{0,1\}^n$ and has entropy $n-k$:

$$D(p_X||u_n) = k$$

subadditivity

if I is a uniform coordinate then X_I is close to uniform:

$$\mathop{\mathbb{E}}_I D(p_{X_i}||u_1) \leq k/n$$

Pinsker

$$\mathop{\mathbb{E}}_I |p_{X_i} - u_1|_1 \leq \sqrt{2k/n}$$

stability?

$X \sim \{0,1\}^n$, $D(p_X||u_n) = k$, $I \sim U([n])$

let $J \sim [n]$ be of high entropy

$$D(p_J||p_I) \leq \epsilon$$

is X_J close to uniform?

stability?

$X \sim \{0,1\}^n$, $D(p_X || u_n) = k$, $I \sim U([n])$

let $J \sim [n]$ be of high entropy

$$D(p_J || p_I) \leq \epsilon$$

is X_J close to uniform?

yes $\mathbb{E}_J \, |p_{X_j} - u_1|_1 \leq |p_J - p_I|_1 + \mathbb{E}_I \, |p_{X_i} - u_1|_1 \leq \sqrt{2\epsilon} + \sqrt{2k/n}$

stability?

$X \sim \{0,1\}^n$, $D(p_X || u_n) = k$, $I \sim U([n])$

let $J \sim [n]$ be of high entropy

$$D(p_J || p_I) \leq \epsilon$$

is X_J close to uniform?

yes $\mathbb{E}_J \, |p_{X_j} - u_1|_1 \leq |p_J - p_I|_1 + \mathbb{E}_I \, |p_{X_i} - u_1|_1 \leq \sqrt{2\epsilon} + \sqrt{2k/n}$

yes $\mathbb{E}_J \, \Delta(p_{X_j}, u_1) \leq 4\epsilon + 10k/n$

proof

$X \sim \{0,1\}^n$, $D(p_X||u_n) = k$, $I \sim U([n])$, $D(p_J||p_I) \leq \epsilon$

for $s \in [n]$ let $g(s) = \Delta(p_{X_s}, u_1)$

write

$$\mathop{\mathbb{E}}_J \Delta(p_{X_j}, u_1) = \mathop{\mathbb{E}}_J g = \mathop{\mathbb{E}}_I g + (\mathop{\mathbb{E}}_J g - \mathop{\mathbb{E}}_I g)$$

the left term

$$\mathop{\mathbb{E}}_I g \leq \mathop{\mathbb{E}}_I 2D(p_{X_i}||u_1) \leq \frac{2k}{n}$$

remains to upper bound the right term $\tau = \mathbb{E}_J g - \mathbb{E}_I g$

proof

$X \sim \{0,1\}^n$, $D(p_X || u_n) = k$, $I \sim U([n])$, $D(p_J || p_I) \leq \epsilon$

for $s \in [n]$ let $g(s) = \Delta(p_{X_s}, u_1)$

upper bound $\tau = \mathbb{E}_J\, g - \mathbb{E}_I\, g$ by

$$|\tau| = \sum_s \frac{(p_J(s) - p_I(s))}{\sqrt{p_J(s) + p_I(s)}} \sqrt{g(s)} \cdot \sqrt{p_J(s) + p_I(s)} \sqrt{g(s)}$$

proof

$X \sim \{0,1\}^n$, $D(p_X \| u_n) = k$, $I \sim U([n])$, $D(p_J \| p_I) \leq \epsilon$

for $s \in [n]$ let $g(s) = \Delta(p_{X_s}, u_1)$

upper bound $\tau = \mathbb{E}_J\, g - \mathbb{E}_I\, g$ by

$$|\tau| = \sum_s \frac{(p_J(s) - p_I(s))}{\sqrt{p_J(s) + p_I(s)}} \sqrt{g(s)} \cdot \sqrt{p_J(s) + p_I(s)} \sqrt{g(s)}$$

$$\leq \sqrt{\sum_s \frac{(p_J(s) - p_I(s))^2}{p_J(s) + p_I(s)} g(s)} \sqrt{\sum_s (p_J(s) + p_I(s)) g(s)}$$

proof

$X \sim \{0,1\}^n$, $D(p_X \| u_n) = k$, $I \sim U([n])$, $D(p_J \| p_I) \leq \epsilon$

for $s \in [n]$ let $g(s) = \Delta(p_{X_s}, u_1)$

upper bound $\tau = \mathbb{E}_J\, g - \mathbb{E}_I\, g$ by

$$
\begin{aligned}
|\tau| &= \sum_s \frac{(p_J(s) - p_I(s))}{\sqrt{p_J(s) + p_I(s)}} \sqrt{g(s)} \cdot \sqrt{p_J(s) + p_I(s)} \sqrt{g(s)} \\
&\leq \sqrt{\sum_s \frac{(p_J(s) - p_I(s))^2}{p_J(s) + p_I(s)} g(s)} \sqrt{\sum_s (p_J(s) + p_I(s)) g(s)} \\
&\leq \sqrt{2 \sum_s \frac{(p_J(s) - p_I(s))^2}{p_J(s) + p_I(s)}} \sqrt{\sum_s (p_J(s) + p_I(s)) g(s)}
\end{aligned}
$$

proof

$X \sim \{0,1\}^n$, $D(p_X \| u_n) = k$, $I \sim U([n])$, $D(p_J \| p_I) \leq \epsilon$

for $s \in [n]$ let $g(s) = \Delta(p_{X_s}, u_1)$

upper bound $\tau = \mathbb{E}_J\, g - \mathbb{E}_I\, g$ by

$$\begin{aligned}
|\tau| &= \sum_s \frac{(p_J(s) - p_I(s))}{\sqrt{p_J(s) + p_I(s)}} \sqrt{g(s)} \cdot \sqrt{p_J(s) + p_I(s)} \sqrt{g(s)} \\
&\leq \sqrt{\sum_s \frac{(p_J(s) - p_I(s))^2}{p_J(s) + p_I(s)} g(s)} \sqrt{\sum_s (p_J(s) + p_I(s)) g(s)} \\
&\leq \sqrt{2 \sum_s \frac{(p_J(s) - p_I(s))^2}{p_J(s) + p_I(s)}} \sqrt{\sum_s (p_J(s) + p_I(s)) g(s)} \\
&= \sqrt{2\Delta(p_J, p_I)} \sqrt{\tau + 2 \mathop{\mathbb{E}}_I g}
\end{aligned}$$

proof

$X \sim \{0,1\}^n$, $D(p_X \| u_n) = k$, $I \sim U([n])$, $D(p_J \| p_I) \le \epsilon$

for $s \in [n]$ let $g(s) = \Delta(p_{X_s}, u_1)$

upper bound $\tau = \mathbb{E}_J\, g - \mathbb{E}_I\, g$ by

$$
\begin{aligned}
|\tau| &= \sum_s \frac{(p_J(s) - p_I(s))}{\sqrt{p_J(s) + p_I(s)}} \sqrt{g(s)} \cdot \sqrt{p_J(s) + p_I(s)} \sqrt{g(s)} \\
&\le \sqrt{\sum_s \frac{(p_J(s) - p_I(s))^2}{p_J(s) + p_I(s)} g(s)} \sqrt{\sum_s (p_J(s) + p_I(s)) g(s)} \\
&\le \sqrt{2 \sum_s \frac{(p_J(s) - p_I(s))^2}{p_J(s) + p_I(s)}} \sqrt{\sum_s (p_J(s) + p_I(s)) g(s)} \\
&= \sqrt{2\Delta(p_J, p_I)} \sqrt{\tau + 2 \mathop{\mathbb{E}}_I g} \le \sqrt{\mathbf{2}\epsilon} \sqrt{\tau + \mathbf{4k/n}}
\end{aligned}
$$

high level

problem: need to analyze $\mathbb{E}_p\, g$ for "complicated" p

possible solution: analyze $\mathbb{E}_q\, g$ for q that is

– "simple"

– Δ-close to p

need to control variances of g

summary

round elimination

lower bound on number of rounds

consider not eliminating

f-divergences

useful collections of “distances”

triangular discrimination

may help to avoid square-root loss