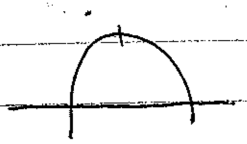
קצת תורת אינפורמציה

עבור $1 \geq p \geq 0$ נגדיר $H(p) = p \cdot \lg\frac{1}{p} + (1-p)\lg\frac{1}{1-p}$

כמה תכונות:
- $H(p) \geq 0$ (מינ')
- פונק' קעורה (נגזרת שנייה)
- מקסימום ב- $\frac{1}{2}$, $H(\frac{1}{2})=1$



באופן כללי: אם A משתנה על $\{1,\dots,k\}$

$$H(A) = \sum_{x=1}^{k} \Pr(A=x) \lg\frac{1}{\Pr(A=x)}$$

פונק' האנטרופיה הבינארית היא עבור המקרה הפרטי של התפלגות על 2 איברים.

תכונות: לכל התפלגות p $H(p) \geq 0$ (מינ')

- אם p מוגדרת על k איברים $H(p) \leq \lg_2 k$

ומקסימום מתקבל כש-p אחידה על k האיברים.

קעורה:

משפט: אם p,q התפלגויות ו- $1 \geq \lambda \geq 0$

אז $H(\lambda p + (1-\lambda) q) \geq \lambda H(p) + (1-\lambda) H(q)$

הוכחה: נגדיר $f(\lambda) = H(\lambda p + (1-\lambda) q)$.

רוצים להראות $f(\lambda) \geq \lambda f(1) + (1-\lambda) f(0)$

ולכן מספיק להראות ש- f קעורה כלומר מספיק $f'' \leq 0$

f במשתנה למד

נסמן:

$$f(\lambda) = \sum_x \left(\lambda p_x + (1-\lambda) q_x\right) \lg \frac{1}{\lambda p_x + (1-\lambda) q_x}$$

$$f'(\lambda) = \sum_x \left[(p_x - q_x) \cdot \lg \frac{1}{\lambda p_x + (1-\lambda) q_x} + \left(\lambda p_x + (1-\lambda) q_x\right) \cdot \frac{(-1)}{c} \cdot \frac{1}{\lambda p_x + (1-\lambda) q_x} (p_x - q_x) \right.$$

לא תלוי λ.

$c>0$ קבוע

$$f''(\lambda) = \sum_x (p_x - q_x) \cdot (-1) \cdot \frac{1}{c} \cdot \frac{1}{\lambda p_x + (1-\lambda) q_x} (p_x - q_x)$$

$(p_x - q_x)^2 \geq 0$, $-1 < 0$, וחיובי ≤ 0

$H(A,B) = H(A) + H(B)$ אז A, B בלתי תלויים

$$\sum_{a,b} \Pr((A,B)=(a,b)) \cdot \lg \frac{1}{\Pr((A,B)=(a,b))}$$

$$= \sum_{a,b} P_a \cdot P_b \cdot \lg \frac{1}{P_a \cdot P_b}$$

$$= \underbrace{\sum_{a,b} P_a P_b \lg \frac{1}{P_a}}_{\sum_a (\sum_b P_b) \cdot P_a \lg \frac{1}{P_a}} + \sum_{a,b} P_a P_b \lg \frac{1}{P_b}$$

$H(A)$ $+$ $H(B)$

משפט כלל:

$$H(A,B) = H(A) + \underbrace{H(B|A)}_{\sum_a H(B|A=a)}$$

הוכחה

$$H(A,B) = \sum_{a,b} P_{a,b} \cdot \lg \frac{1}{P_{a,b}}$$

$$= \sum_{a,b} P_a \cdot P_{b|a} \cdot \lg \left(\frac{1}{P_a} \cdot \frac{1}{P_{b|a}} \right)$$

$$= \underbrace{\sum_a \left(\sum_b P_{b|a} \right) \cdot P_a \cdot \lg \frac{1}{P_a}}_{H(A)} + \underbrace{\sum_a P_a \cdot \sum_b P_{b|a} \cdot \lg \frac{1}{P_{b|a}}}_{H(B|A=a)}$$

$H(B|A)$

מסקנות: 1. $H(B|A) \geq 0$ (שוויון אם B פונקציה של A)

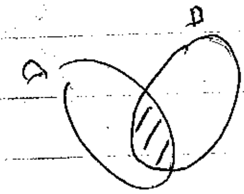
2. $H(A) \leq H(A,B)$ $\forall B$

כי:

$$H(B) = H\left(\sum_a p_a \cdot (B|A=a)\right)$$

$$\underset{\text{קמירות}}{\geq} \sum_a p_a H(B|A=a) = H(B|A)$$

4. $H(A,B) = H(A) + H(B|A) \leq H(A) + H(B)$



סימון: $I(A:B) = H(A)+H(B)-H(A,B)$

האינפורמציה ההדדית של A, B.

תכונות: $\forall_{A,B} \quad I(A;B) \geq 0$

דוגמאות: 1. אם A,B בת"ל אז $I(A:B) = H(A)+H(B)-(H(A)+H(B)) = 0$

2. אם (A,B) מקבל ערכים (a,a) אז

$$I(A:B) = H(A) + H(A) - \underset{H(A)}{H(A,B)} = H(A)$$

כי B לא מוסיף מידע על A.

$$H(p) = \underbrace{\sum_x p_x \lg\frac{1}{p_x}}_{\text{כמות ה"הפתעה" של X הצפויה}} \qquad \underbrace{\sum_x p_x \lg\frac{1}{q_x}}_{\text{כמות ההפתעה כשאנחנו מאמינים לפי q}}$$

ולא חושבים שמאמינים לפי q.

ההפרש: $D(P\|Q) = \sum_x p_x\left(\lg\frac{1}{q_x} - \lg\frac{1}{p_x}\right) = \sum_x p_x \cdot \lg\frac{p_x}{q_x}$

(4) משפט: לכל P,Q $D(P\|Q) \geq 0$

הוכחה: $D(P\|Q) = \sum_x p_x \cdot -\lg\left(\frac{q_x}{p_x}\right) \leq -\log\left(\sum_x p_x \cdot \frac{q_x}{p_x}\right) = \lg\left(\sum_x q_x\right) = \lg(1) = 0$

__בעיה:__ Alice ו Bob מחזיקים קלט מהתפלגות משותפת __ידועה__ (A,B). Alice מחזיקה a וב' מחזיק b.

Alice ו Bob צריכים פרוטוקול תקשורת (כלומר כיוון אחד מספר ביטים) שבסופו גם Alice וגם Bob יודעים את a.

כמה ביטים של תקשורת נדרשים בפרוטוקול?

__פתרון:__ לפחות $H(A|B)$

~~[illegible]~~

בתחילת התהליך האינפו' ההדדית של Alice ו Bob היא $I(A:B)$
בסוף " " " " " " היא $H(A)$

כל מסר C שעובר, יכול להגדיל את האינפו' ההדדית
לכל היותר ב $H(C)$, כי:

$$
\begin{aligned}
I(XC:Y) &= H(XC) + H(Y) - H(XCY) \\
&= H(X) + H(C|X) + H(Y) - \left(H(XY) + H(C|XY) \right) \\
&= I(X:Y) + H(C|X) - H(C|XY) \\
&\overset{H\geq 0}{\leq} I(X:Y) + H(C|X) \\
&\leq I(X:Y) + H(C)
\end{aligned}
$$

ולכן, אם עברו המסרים $C_1 \ldots C_m$, $|C_i| = l_i$

$$
\begin{aligned}
H(A) &\leq I(A:B) + H(C_1) + H(C_2) \ldots + H(C_m) \\
&\leq I(A:B) + l_1 + \ldots + l_m
\end{aligned}
$$

ולכן $l = l_1 + \ldots + l_m \geq H(A) - I(A:B)$
$= H(AB) - H(B)$
$= H(A|B)$.

Quantum information theory

עבור מצב קוונטי ρ_A

מתקיים $Tr(\rho)=1$, $\rho \geq 0$

לכן $\rho = \sum \lambda_i |\psi_i\rangle\langle\psi_i|$ עבור איזה בסיס חלון

ואם נלכסן $\rho = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix}$ בבסיס הזה

נגדיר $H(\rho) = H(\lambda_1 \ldots \lambda_N)$ (הגדרה שקולה לסימון $V(\rho)$

(Von Neumann entropy השם

תכונות:

1. לכל ρ $H(\rho) \geq 0$ ברור.

2. אם ρ מוגדר על מרחב הילברט ממימד N

אז $H(\rho) \leq \log_2(N)$. ברור.

מתקיים ל- completely mixed state $\frac{1}{N}I$. ברור.

3. עבור pure state $H(\rho)=0 \Leftrightarrow \rho=|\psi\rangle\langle\psi|$

(נראה גם תכונות של H שבורה, למרות שצריך לעבוד יותר קשה.

ואפשר:

אי שוויון $H(A,B) \leq H(A)+H(B)$

כלומר לכל $\rho_{A,B}$ נסמן $\rho_A = Tr_B(\rho_{AB})$, $\rho_B = Tr_A(\rho_{AB})$

אז $H(\rho_{AB}) \leq H(\rho_A) + H(\rho_B)$

תזכורת:

$$S(\rho) = -\mathrm{Tr}(\rho \lg \rho)$$

נגדיר

$$D(\rho \| \sigma) = \mathrm{Tr}(\rho(\lg \rho - \lg \sigma))$$

משפט (**Klein**): לכל ρ, σ, $D(\rho\|\sigma) \geq 0$.

הוכחה

$\sigma = \sum q_i w_i w_i^*$

$\rho = \sum p_i v_i v_i^*$

נניח v_i הם ו"ע של ρ

w_i " (ו"ע) " σ

$$D(\rho\|\sigma) = \mathrm{Tr}\Big(\sum_i p_i \lg p_i \, v_i v_i^* - \sum_{i,j} p_i \log q_j \, v_i v_i^* w_j w_j^* \Big)$$

$$= \sum_i p_i \log p_i - \sum_{i,j} p_i \log q_j \, |v_i^* w_j|^2$$

$$= \sum_i p_i \Big[\log p_i - \sum_j P_{ij} \log q_j \Big]$$

כאשר $P_{ij} = |v_i^* w_j|^2$, ו- $P = (P_{ij})$ doubly stochastic

כי:

$$\sum_i P_{ij} = \sum_i |v_i^* w_j|^2 = \|w_j\| = 1$$

$$\sum_j P_{ij} = \sum_j |v_i^* w_j|^2 = \|v_i\| = 1$$

$$\geq \sum_i p_i \Big[\log p_i - \log\Big(\sum_j P_{ij} q_j \Big) \Big]$$

$-\log$ קמורה

נסמן $r_i = \sum_j P_{ij} q_j$ אז $r_i \geq 0$

$$\sum_i r_i = \sum_{i,j} P_{ij} q_j = \sum_j q_j \underbrace{\sum_i P_{ij}}_{1} = \sum_j q_j = 1$$

ולכן r התפלגות

$$= D(p \| r) \geq 0$$

(מהמשפט הקודם!) $\square$

משפט: לכל ρ_{AB} : $H(\rho_{AB}) \leq H(\rho_A) + H(\rho_B)$

הוכחה: נבחר $\sigma_{AB} = \rho_A \otimes \rho_B$

אז $H(\rho_{AB}) = -Tr(\rho \lg \rho) \leq -Tr(\rho \lg \sigma)$ (Klein משפט)

$= -Tr\left(\rho_{AB} \left(\log (\rho_A \otimes I)(I \otimes \rho_B)\right)\right)$ (מתחלפים)

$= -Tr_{A,B}\left(\rho_{AB} \log(\rho_A \otimes I)\right) - Tr\left(\rho_{AB} \log (I \otimes \rho_B)\right)$, where $\log(\rho_A \otimes I) = (\log \rho_A) \otimes I$

$= -Tr_A(\rho_A \log \rho_A) - Tr(\rho_B \log \rho_B)$

$= H(A) + H(B)$

לכן $I(A:B) = H(\rho_{AB}) - H(\rho_A) - H(\rho_B)$, $I(A:B) \geq 0$

משפט: H קמורה

$$H\left(\sum \lambda_i \rho_i\right) \geq \sum \lambda_i H(\rho_i)$$

הוכחה: נניח $\rho_A = \sum \lambda_i \rho_i$

נסמן $\rho_{AB} = \sum \lambda_i \underbrace{\rho_i}_{A} \otimes \underbrace{|i\rangle\langle i|}_{B}$

$S(A) = S(\rho)$

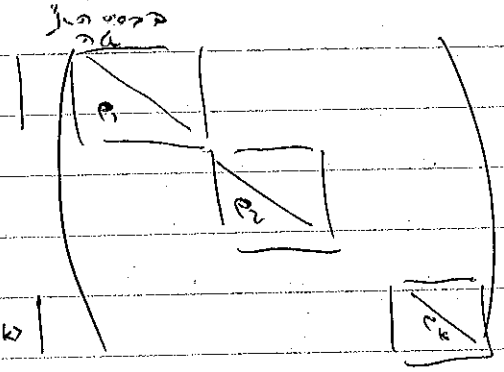
$S(B) = S\left(\sum \lambda_i |i\rangle\langle i| \, Tr(\rho_i)\right) = H(\lambda_1 \ldots \lambda_N) = H(\lambda)$, where $Tr(\rho_i) = 1$

(הסבר במשאל קודם) $S(A,B) = H(\lambda) + \sum_i \lambda_i S(\rho_i)$

$\Leftarrow$ $S(A,B) \leq S(A) + S(B)$

$$\boxed{\sum \lambda_i S(\rho_i) \leq S(\rho)}$$ □

$$S\left(\sum_{i=1}^{k} \lambda_i \rho_i \otimes |i\rangle\langle i|\right) =$$



$$= \sum_{i=1}^{k} \sum_{j=1}^{N} \lambda_i p_{ij} \lg \frac{1}{\lambda_i p_{ij}}$$

$$= \sum_{i=1}^{k} \sum_{j=1}^{N} \lambda_i p_{ij} \cdot \lg \frac{1}{\lambda_i} + \sum_i \lambda_i \sum_j p_{ij} \lg \frac{1}{p_{ij}}$$

$$H(\lambda) \qquad + \sum_i \lambda_i S(\rho_i)$$

משפט Holevo:

נניח A מעביר $x \in X$, X מפולג לפי p_x.

נניח B מקבל קידוד קוונטי של x, $\rho(x) \in H$.

אז לכל מדידה ש-B יבצע, והאינפורמציה הקלאסית שהוא יקבל:

$$I(A:B) \le \underbrace{S\left(\sum_x p_x \rho_x\right)}_{\text{Shannon ent של קידוד ממוצע}} - \sum_x p_x S(\rho_x)$$

הערה: אם ρ_x הוא pure state אז:

$$I(A:B) \le S\left(\sum_x p_x \rho_x\right)$$

(הערה) אם עובדים $\rho \in H_K$, ערוץ הקידוד ממימד $K = 2^k$:

$$I(A:B) \le k = \log(K)$$

במילים: אי אפשר להעביר יותר מ-k ביטים של אינפורמציה ע"י k ביטים קוונטים.

בקוונטים זה בהכרח לא נכון ש $H(A) \le H(A,B)$

דוגמא: $\rho_{AB} = |\Psi_{AB}\rangle\langle\Psi_{AB}|$ עבור $\Psi_{AB} = \frac{1}{\sqrt{2}}\left[|00\rangle + |11\rangle\right]$

לכן $H(\rho_{AB}) = 0$,

אבל: $\rho_B = \rho_A = \begin{pmatrix} 1 & \\ & 1 \end{pmatrix}$ הם $H(A) = H(B) = 1$,

RAC

A משייכת $x \in_R \{0,1\}^n$

מחפשים קבוצה $p(x) \in H_k$ עבור k כמה שיותר קטן.

רק פעם תחזיר 1

יש מסבה: Θ_i, שמסתדרת $\leq p$ תחזיר את x_i

עכשיו נראה שאי אפשר להגיע שמשתמשים נחזיר את כל $x_1 ... x_n$

כי $I(A:B) \leq lg(k)=k$

אבל יתכן שנוכל לעשות קידוד רק אחד שיחזיר רק בC

[illegible] ... C

[illegible]

נראה שזה לא מתאים:

טענה: $k \geq (1-H(p)) \cdot m$

הוכחה: נסמן $X_1 .. X_n$ [illegible]

C [illegible]

טענה $I(X_1X_2:C) = I(X_1:C) + I(X_2:X_1C)$

כאשר X_1, X_2 ב"ת

הוכחה: $H(X_1X_2) + H(C) - H(X_1X_2X_1) \overset{?}{=} H(X_1) + H(C) - H(X_1C) + H(X_2) + H(X_1C) - H(X_1X_2C)$

$H(X_1X_2) = H(X_1) + H(X_2)$

למסקנה כאשר $X_1 \ldots X_n$ ק"מ

$$I(X_1 \ldots X_n : c) = I(X_1 : c) + I(X_2 : X_1 c) + \ldots + I(X_n : X_1 \ldots X_{n-1} c)$$

מסקנה אם c ניתן לניחוש ע"י X_i בהסתברות $\geq p$

אז $I(X_i : c) \geq (1-H(p))c$.

הוכחה

$$I(X_1 \ldots X_n c) = \sum_{i=1}^{n} I(X_i : c X_1 \ldots X_{i-1})$$

$$\geq \sum_{i=1}^{n} I(X_i : c)$$

$$\boxed{\geq n(1-H(p))}$$

תרגילי חזרה

① יהיו משתנים מקריים X,Y אשר $Pr(X=Y)=p$, X,Y בלתי תלויים, X אחיד

אז $I(X:Y) \geq 1-H(p)$.

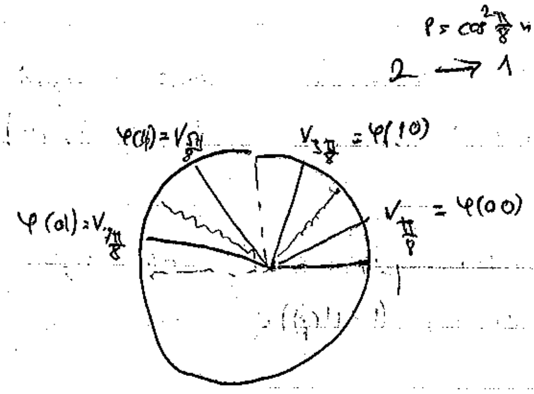
[illegible]

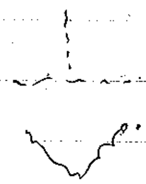
② יהיו X,Y קבועים,

Y' קלט, תוצאה הפעלה שגבה על Y

אז $I(X:Y') \leq I(X:Y)$.

RAC ל [illegible]





תרגילים: 1. RAC ל $2 \overset{0.86}{\to} 1$

2. להזכיר $R''(EQ)$ ידוע כשבכה $R''(EQ) = \Theta(\sqrt{n})$

$Q''(EQ)$

להראות $Q''(EQ) = O(\lg n)$

3. להוכיח את למה ⑩ מקודם.