情報理論　#4

第4回講義　5月16日

情報科学研究科　　　井上純一

http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/

Kulback-Leibler情報量

$$\underline{D(P \| Q)} = \sum_{x \in X} P(x) \underline{\{\log P(x) - \log Q(x)\}}$$

KL情報量

2つの確率分布P,Q間の距離

$$= \sum_{x \in X} P(x) \log \left\{ \frac{P(x)}{Q(x)} \right\}$$

2確率変数からなる分布 $P_{XY}(x, y), Q_{XY}(x, y)$ に対しても同様に

$$D(P \| Q) = \sum_{x \in X} \sum_{y \in Y} P_{XY}(x, y) \{\log P_{XY}(x, y) - \log Q_{XY}(x, y)\}$$

$$= \sum_{x \in X} \sum_{y \in Y} P_{XY}(x, y) \log \left\{ \frac{P_{XY}(x, y)}{Q_{XY}(x, y)} \right\}$$

で定義される

KL情報量と相互情報量

同時分布に関する周辺分布を $P_X(x)=\sum_{y\in Y} P_{XY}(x,y)$, $P_Y(y)=\sum_{x\in X} P_{XY}(x,y)$

で導入すると、確率変数X、Yの間の相互情報量は

$$
\begin{aligned}
I(X;Y) &= -\sum_{x\in X} P_X(x)\log P_X(x) - \sum_{y\in Y} P_Y(y)\log P_Y(y) + \sum_{x,y\in X,Y} P_{XY}(x,y)\log P_{XY}(x,y) \\
&= -\sum_{x,y\in X,Y} P_{XY}(x,y)\log P_X(x) - \sum_{x,y\in X,Y} P_{XY}(x,y)\log P_Y(y) + \sum_{x,y\in X,Y} P_{XY}(x,y)\log P_{XY}(x,y) \\
&= \sum_{x,y\in X,Y} P_{XY}(x,y)\log\left\{\frac{P_{XY}(x,y)}{P_X(x)P_Y(y)}\right\} = \underline{D(P_{XY}\,\|\,P_X P_Y)}
\end{aligned}
$$

確率変数X、Yが独立な場合最小（ゼロ）となる。

$I(X;Y)=H(X)-H(X\mid Y)$　より、Yを知ったときに

Xに関する知識の増加分はゼロ

イェンセンの不等式

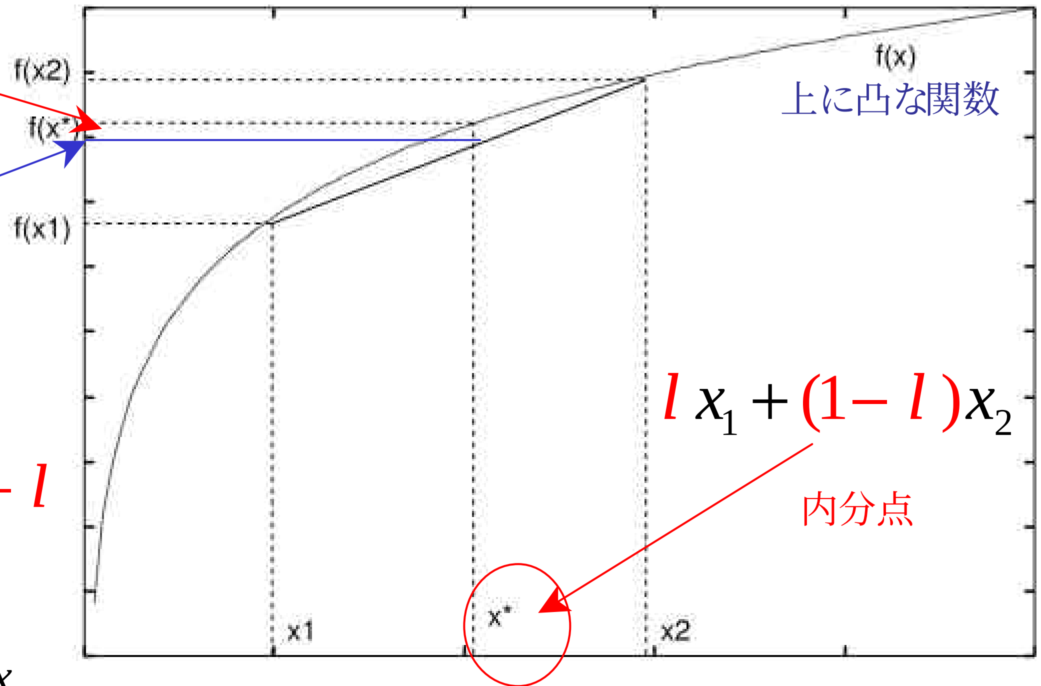
$f(\lambda x_1+(1-\lambda)x_2)$

内分点での関数値

常に成り立つ

$\geq \lambda f(x_1)+(1-\lambda)f(x_2)$

関数値の内分点



$\lambda x_1+(1-\lambda)x_2$

内分点

$$P(x=x_1)=\lambda,\quad P(x=x_2)=1-\lambda$$

とすると

$$E(x)=\sum_{x=x_1,x_2} xP(x)=\lambda x_1+(1-\lambda)x_2$$

$$f(E(x))=f(\lambda x_1+(1-\lambda)x_2)$$

$$E(f(x))=\lambda f(x_1)+(1-\lambda)f(x_2)$$

$$\boxed{f(E(x))\geq E(f(x))}$$

イェンセンの不等式

KL情報量の非負性

イェンセンの不等式で　$f(x) \mapsto \log x,\ \ x \mapsto Q(x)/P(x)$　とすると

$$
\begin{aligned}
-D(P \| Q) &= \sum_{x \in X} P(x) \log \left\{ \frac{Q(x)}{P(x)} \right\} \\
&= E\left(f\left(\frac{Q(x)}{P(x)} \right) \right) \\
&\le f\left(E\left(\frac{Q(x)}{P(x)} \right) \right) = \log \left\{ \sum_{x \in X} P(x) \cdot \frac{Q(x)}{P(x)} \right\} = \log 1 = 0
\end{aligned}
$$

$$\boxed{D(P \| Q) \ge 0}$$

KL情報量は負にならない

情報源圧縮

Ababbabaabbb という系列を ab|ab|ba|ba|aa|bb と2つずつに区切り、
次の表に従って01の並びからなる符号に変換する

アルファベット	符号語	符号長
aa	110	3
ab	0	1
ba	10	2
bb	111	3

この表による変換で

ababbabaabbb 12ビット

→ 0010100111 10ビット

に圧縮される

対応表を逆にたどると元の系列
が復元できるので**可逆圧縮**である
復号

どの符号語も他の符号語の先頭部分になっていない
⇒ **語頭条件**がみたされている

いくつかの概念と定義

$A=\{a_1,a_2,..,a_M\}$ ：情報源をつくる記号の集合

$B=\{b_1,b_2,...,b_K\}$ ：通信路で使われる記号の集合

符号化 : Aからなる系列をBからなる系列に変換すること

復号化 : Bからなる系列から、元のAからなる系列を再生すること

$x_i \in B \ (i=1,..,K)$ **から符号語** $x=x_1\cdots x_k$ **を作った場合、**k **を符号長と呼ぶ**

$$c^+=\{0,1,00,01,10,11,000,.....\},\ c=\{0,1\}$$

$$c^n=\{x_1x_2\cdots x_n : x_i\in c,\ i=1,...,n\}$$

復号可能条件

復号できるための条件

$A \subset \mathcal{A}^n, B \subset \mathcal{B}^+$とし、符号化は写像：

$y: A \to B$で行われるものとすると、$\forall_{x,x'} \in A$に対し

$$x \neq x' \quad \Rightarrow \quad y(x) \neq y(x')$$

のとき、符号yは正則であると言う。また

$$x \neq x' \quad \Leftrightarrow \quad y(x) \neq y(x')$$

が成り立つとき、この符号yは一意復号可能である

※ 語頭条件を満たしている正則な符号は一意復号可能である。

クラフト不等式

$f: A \to B^{+}$は一意復号可能であるとし、$l(x)$をxに対する符号語$f(x)$の長さとすると

$$\sum_{x \in A} K^{-l(x)} \le 1$$

が成り立つ。　KはBの要素数である。

適用例）

$$A = \{aa, ab, ba, bb\}, B = \{0,1\}$$

$$y_2(aa) = 00, y_2(ab) = 10, y_2(ba) = 11, y_2(bb) = 110$$

$$l(aa) = l(ab) = l(ba) = 2, l(bb) = 3, K = 2$$

のときは

$$\sum_{x \in A} K^{-l(x)} = 2^{-2} + 2^{-2} + 2^{-2} + 2^{-3} = \frac{7}{8} < 1$$

一意復号可能

クラフト不等式の証明 #1

J を任意の正の整数として、次が成り立つことに注意する

$$\left(\sum_{x\in A} K^{-l(x)}\right)^J = \boxed{\sum_{x_1\in A}\sum_{x_2\in A}\cdots\sum_{x_J\in A} K^{-(l(x_1)+l(x_2)+\cdots l(x_J))}}$$

$A_J(n)$ を全体の長さが n になるようなJ 個の符号語の並べ方の総数とすると

$$A_J(n) = \#\left\{x_1x_2\cdots x_J \in A^J :\ l(x_1)+l(x_2)+\cdots l(x_J)=n\right\}$$

であり、ある符号語に対し、次が成り立つ

J個の符号語を並べた符号長の最大値

$$\boxed{\sum_{n=1}^{Jl_{\max}} d\left(l(x_1)+\cdots+l(x_J),n\right)=1}$$

1をかける

クラフトの不等式の証明 #2

$$\left(\sum_{x\in A} K^{-l(x)}\right)^J = \sum_{n=1}^{Jl_{\max}} \sum_{x_1\in A} \cdots \sum_{x_J\in A} d\left(l(x_1)+\cdots+l(x_J), n\right) K^{-n} = \sum_{n=1}^{Jl_{\max}} A_J(n) K^{-n}$$

$\forall_{a,a'} \in A$**に対して** $f(a) \neq f(a')$**となるためには**

$$A_J(n) \leq K^n$$

Bの記号をn個並べてできる語の総数

であることが必要。

よって
$$\left(\sum_{x\in A} K^{-l(x)}\right)^J = \sum_{n=1}^{Jl_{\max}} A_J(n) K^{-n} \leq \sum_{n=1}^{Jl_{\max}} 1 = Jl_{\max}$$

$$\longrightarrow \quad \sum_{x\in A} K^{-l(x)} \leq \left(Jl_{\max}\right)^{1/J} = 1 \; (J \to \infty)$$

クラフト不等式を用いた語頭符号の構成

$B=\{0,1\}\,(K=2), l_1=1, l_2=2, l_3=l_4=3$　とすると

$2^{-1}+2^{-2}+2^{-3}+2^{-3}=1\leq 1$　となりクラフト不等式をぎりぎりで満たす。

符号の木による表現