



HOKKAIDO UNIVERSITY

Title	2004年度 情報理論講義ノート
Author(s)	井上, 純一; Inoue, Jun-ichi
Description	この講義資料は著者のホームページ http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/ からもダウンロードできます。 http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/
Issue Date	2004
Doc URL	https://hdl.handle.net/2115/374
Rights(URL)	https://creativecommons.org/licenses/by-nc-sa/2.1/jp/
Type	learning object
File Information	InfoTheory04_4.pdf, 第4回講義ノート



情報理論 配布資料 #4

担当：井上 純一 (情報エレクトロニクス系棟 8-13)

平成 16 年 5 月 17 日

演習問題 2 に関するコメント

まず, 前回配布した解答に一部間違いがありました. 問題 1. の (2) で相互情報量を求める際, 条件付きエントロピー $H(A|B)$ の計算に間違いがあります. 解答の (12) 式にそれぞれの確率を代入して行けばよいわけですが, 出力値の確率 $p(b_1), p(b_2)$ のそれぞれが

$$\begin{aligned} p(b_1) &= \sum_{j=1,2} p(b_1|a_j)p(a_j) \\ &= p(b_1|a_1)p(a_1) + p(b_1|a_2)p(a_2) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} \end{aligned} \quad (1)$$

$$\begin{aligned} p(b_2) &= \sum_{j=1,2} p(b_2|a_j)p(a_j) \\ &= p(b_2|a_1)p(a_1) + p(b_2|a_2)p(a_2) = \frac{1}{2} \times \frac{1}{2} + 1 \times \frac{1}{2} = \frac{3}{4} = 1 - p(b_1) \end{aligned} \quad (2)$$

で与えられますから, 条件付きエントロピー $H(A|B)$ は

$$H(A|B) = -\frac{1}{3} \times \frac{3}{4} \log \frac{1}{3} - \frac{2}{3} \times \frac{3}{4} \log \frac{2}{3} = \frac{3}{4} \log 3 - \frac{1}{2} \quad (3)$$

となります. 従って, 求める相互情報量 $I(A; B)$ は

$$\begin{aligned} I(A; B) &= H(A) - H(A|B) \\ &= 1 - \left\{ \frac{3}{4} \log 3 - \frac{1}{2} \right\} = \frac{3}{2} - \frac{3}{4} \log 3 \simeq 0.31 \text{ (ビット)} \end{aligned} \quad (4)$$

と計算できます. これが正しい答えです.

また, 問題 2. では「手付かず」の人が意外に多かったのですが, ラグランジュの未定係数法という手法は案外使う場面が多いと思いますので, 配布の解答を読んで確認しておいてください.

演習問題 3 の解答例

(1) 定常的無記憶情報源の場合, 結合確率は

$$p(x_1, x_2, \dots, x_n) = \prod_{i=1}^n p(x_i) \quad (5)$$

と書け、この中の確率 $p(x_i)$ のそれぞれが

$$p(x_i) = p\delta_{x_i,1} + (1-p)\delta_{x_i,0} \quad (6)$$

で与えられるので、この系のエントロピーは

$$\begin{aligned} H(x_1, x_2, \dots, x_n) &= -\sum_{x_1} \sum_{x_2} \cdots \sum_{x_n} \prod_{i=1}^n p(x_i) \log \prod_{i=1}^n p(x_i) \\ &= -\sum_{x_1} \sum_{x_2} \cdots \sum_{x_n} p(x_1)p(x_2)\cdots p(x_n) \{\log p(x_1) + \cdots + \log p(x_n)\} \\ &= -n \sum_{x_1=0,1} p(x_1) \log p(x_1) \sum_{x_2} \sum_{x_3} \cdots \sum_{x_n} p(x_2)\cdots p(x_n) \\ &= -n \sum_{x_1=0,1} \{p\delta_{x,1} + (1-p)\delta_{x,0}\} \log \{p\delta_{x,1} + (1-p)\delta_{x,0}\} \\ &= -n\{(1-p)\log(1-p) + p\log p\} \end{aligned} \quad (7)$$

であるから、エントロピーレート \mathcal{H} は

$$\mathcal{H} = \lim_{n \rightarrow \infty} \frac{1}{n} H(x_1, x_2, \dots, x_n) = -p\log p - (1-p)\log(1-p) \quad (8)$$

となる。

(2) まず、この定常的マルコフ情報源の状態遷移をグラフで表すと図1のようになる。この場合のエントロ

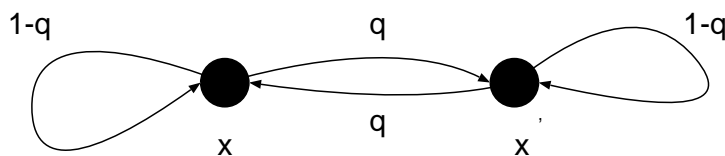


図1: 遷移確率 $p(x'|x) = q + (1-2q)\delta_{x',x}$ のグラフ表現.

ピーは結合確率が

$$p(x_1, x_2, \dots, x_n) = p(x_n|x_{n-1})p(x_{n-1}|x_{n-2})\cdots p(x_3|x_2)p(x_2|x_1)p(x_1) \quad (9)$$

と書けることに注意すれば、この系のエントロピーは

$$\begin{aligned} H(x_1, x_2, \dots, x_n) &= -\sum_{x_1} \cdots \sum_{x_n} p(x_1, \dots, x_n) \{\log p(x_1) + \log p(x_2|x_1) + \cdots + \log p(x_n|x_{n-1})\} \\ &= -\sum_{x_1} p(x_1) \log p(x_1) - \sum_{x_1} \sum_{x_2} p(x_1, x_2) \log p(x_2|x_1) - \cdots \\ &\quad - \sum_{x_n} \sum_{x_{n-1}} p(x_n, x_{n-1}) \log p(x_n|x_{n-1}) \\ &= -\sum_{i=1}^n \left\{ \sum_{x_{i-1}} p(x_{i-1}) \sum_{x_i} p(x_i|x_{i-1}) \log p(x_i|x_{i-1}) \right\} \\ &= -n \sum_x p(x) \sum_{x'} p(x'|x) \log p(x'|x) \end{aligned} \quad (10)$$

で与えられる。従って、エントロピーレート \mathcal{H} は

$$\mathcal{H} = \lim_{n \rightarrow \infty} \frac{1}{n} H(x_1, \dots, x_n) = - \sum_x p(x) \sum_{x'} p(x'|x) \log p(x'|x) \quad (11)$$

となる。ここに問題文に与えられた $p(x), p(x'|x)$ を代入し、 x, x' に関する和を実行すれば

$$\begin{aligned} \mathcal{H} &= - \sum_{x=0,1} \{p \delta_{x,0} + (1-p) \delta_{x,1}\} \sum_{x'=0,1} \{q + (1-2q) \delta_{x,x'}\} \log\{q + (1-2q) \delta_{x,x'}\} \\ &= -p \sum_{x'=0,1} \{q + (1-2q) \delta_{0,x'}\} \log\{q + (1-2q) \delta_{0,x'}\} \\ &\quad - (1-p) \sum_{x'=0,1} \{q + (1-2q) \delta_{1,x'}\} \log\{q + (1-2q) \delta_{1,x'}\} \\ &= -p\{(1-q) \log(1-q) + q \log q\} - (1-p)\{q \log q + (1-q) \log(1-q)\} \\ &= -q \log q - (1-q) \log(1-q) \end{aligned} \quad (12)$$

となる。

演習問題 4

1. $K = 3, \mathcal{B} = \{0, 1, 2\}, l_1 = 1, l_2 = 2, l_3 = l_4 = 3$ の符号語に関して以下の問いに答えよ.

- (1) この符号は一意復号可能かどうかを判定せよ (そのように判定した理由も明記すること).
- (2) (1) で判定した結果, 一意復号可能である場合, 語頭条件を満たすような符号語 x_1, x_2, x_3, x_4 を一つ挙げよ.
- (3) 教科書 p. 35 の図 3.4 にならって, この符号語の符号の木を描け.

2. 次の表に与えた符号 ψ_2 (教科書 p.31 参照) について以下の問いに答えよ.

aa	00
ab	10
ba	11
bb	110

- (1) 情報源の生成確率が $p(\text{aa}) = p(\text{ab}) = p(\text{ba}) = p(\text{bb}) = 1/4$ のとき, 情報源のエントロピー H , 及び, 平均符号長 L を求めよ.
- (2) 情報源の生成確率を $p(\text{aa}) = p(\text{ab}) = p(\text{ba}) = p$, 及び, $p(\text{bb}) = 1 - 3p$ と選ぶとき

$$\Phi(p) \equiv L - H \tag{13}$$

を求め, Φ を最小にするような $p = p_*$ 及び, そのときの Φ の値 $\Phi(p_*)$ を求めよ.

- (3) (1)(2) の結果から, 符号 ψ_2 に関してわかることを簡潔に述べよ.