情報理論 配布資料 #3

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 10 日

演習問題 1 に関するコメント

皆さんのレポートを採点してみると, 問題 (4) で確率 $p(k,l)$ を導く際に

$$\begin{array}{rcl} 10 \text{ ビットの中の } k \text{ ビットが消失する確率} & : & {}_{10}C_k\, q^k\, (1-q)^{10-k} \\ \text{残りの } 10-k \text{ ビットの中の } l \text{ ビットが反転する確率} & : & {}_{10-k}C_l\, p^l\, (1-p)^{10-k-l} \end{array}$$

とし, 求める確率をこの両者の積：

$$p(k,l) \quad = \quad {}_{10}C_k\, q^k\, (1-q)^{10-k} \times {}_{10-k}C_l\, p^l\, (1-p)^{10-k-l} \tag{1}$$

としている解答が数多く見受けられました. 一見するとこれで良さそうなのですが, 誤りです.

ここに出てくる 2 番目の確率「残り $10-k$ ビットの中の l ビットが反転する確率」ですが, 正しくは下記のような条件付き確率で与えられます.

$${}_{10-k}C_l \times p(\textbf{反転} \mid \textbf{消失せず})^l \times p(\textbf{反転せず} \mid \textbf{消失せず})^{10-k-l} \tag{2}$$

これら 2 つの条件付き確率は積の公式から

$$p(\textbf{反転} \mid \textbf{消失せず}) \quad = \quad \frac{p(\textbf{反転}, \textbf{消失せず})}{p(\textbf{消失せず})} = \frac{p}{1-q} \tag{3}$$

$$p(\textbf{反転ぜず} \mid \textbf{消失せず}) \quad = \quad \frac{p(\textbf{反転なし}, \textbf{消失せず})}{p(\textbf{消失せず})} = \frac{1-p-q}{1-q} \tag{4}$$

となりますから, (2) 式は

$${}_{10-k}C_l \left\{ \frac{p}{1-q} \right\}^l \left\{ \frac{1-p-q}{1-q} \right\}^{10-k-l} \tag{5}$$

と書き換えることができます. 正しい確率 $p(k,l)$ はこれと ${}_{10}C_k\, q^k\, (1-q)^{10-k}$ の積で与えられるべきであり

$$\begin{aligned} p(k,l) \quad &= \quad {}_{10}C_k\, q^k\, (1-q)^{10-k} \times {}_{10-k}C_l \left\{ \frac{p}{1-q} \right\}^l \left\{ \frac{1-p-q}{1-q} \right\}^{10-k-l} \\ &= \quad {}_{10}C_k\ {}_{10-k}C_l\, q^k\, p^l\, (1-p-q)^{10-k-l} \end{aligned} \tag{6}$$

となります. もちろん, これは前回配布の解答例にある $p(k,l)$ と一致します.

ところで, 続く問題 (5) で $\sum_k \sum_l p(k,l)$ を計算する際, 間違った確率 (1) を用いても

$$\begin{aligned} p_{\rm total} = \sum_k \sum_l p(k,l) &= \sum_{k=0}^{10} {}_{10}C_k\, q^k\,(1-q)^{10-k} \sum_{l=0}^{10-k} {}_{10-k}C_l\, p^l\,(1-p)^{10-k-l} \\ &= \sum_{k=0}^{10} {}_{10}C_k\, q^k\,(1-q)^{10-k}\{p+1-p\}^{10-k} = (q+1-q)^{10} = 1 \end{aligned} \tag{7}$$

となり, 偶然正解と一致します. これはもちろん正しくないのですが, 今回は問題 (5) を独立した小問とみなし, このようなプロセスで $p_{\rm total}=1$ が得られても「結果オーライ」で丸にしました.

演習問題 2 の解答例

1. 問題文の誘導に従って相互情報量を求める.

(1) まず

$$\begin{aligned} \sum_{j=1,2} p(b_j|a_i) &= \sum_{j=1,2} \frac{p(b_j,a_i)}{p(a_i)} \\ &= \frac{\sum_{j=1,2} p(a_i|b_j)p(b_j)}{p(a_i)} = \frac{p(a_i)}{p(a_i)} = 1 \end{aligned} \tag{8}$$

となり, 与えられた関係式の成立が示される. 従って, これを用いれば

$$p(b_2|a_1) = 1 - p(b_1|a_1) = \frac{1}{2} \tag{9}$$

$$p(b_1|a_2) = 1 - p(b_2|a_2) = 0 \tag{10}$$

が直ちに得られる.

(2) エントロピー $H(\mathcal{A})$, 条件付きエントロピー $H(\mathcal{A}|\mathcal{B})$ を計算しよう.

$$H(\mathcal{A}) = -\sum_{j=1,2} p(a_j)\log p(a_j) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = 1 \tag{11}$$

$$\begin{aligned} H(\mathcal{A}|\mathcal{B}) &= -\sum_{i=1,2}\sum_{j=1,2} p(a_i,b_j)\log p(a_i|b_j) \\ &= -\sum_{i=1,2}\sum_{j=1,2} p(a_i|b_j)p(b_j)\log p(a_i|b_j) \\ &= -p(a_1|b_1)p(b_1)\log p(a_1|b_1) - p(a_2|b_1)p(b_1)\log p(a_2|b_1) \\ &- \ p(a_1|b_2)p(b_2)\log p(a_1|b_2) - p(a_2|b_2)p(b_2)\log p(a_2|b_2) \end{aligned} \tag{12}$$

さて, ここから先の計算を進めるためには, 確率 $p(a_1|b_2), p(a_1|b_2)$ 等が必要であるが, これは問題文に与えたヒントの公式より

$$p(a_1|b_1) = \frac{p(b_1|a_1)p(a_1)}{\sum_{j=1,2} p(b_1|a_j)p(a_j)} = \frac{p(b_1|a_1)p(a_1)}{p(b_1|a_1p(a_1) + p(b_1|a_2)p(a_2)} = 1 \tag{13}$$

$$p(a_2|b_1) = \frac{p(b_1|a_2)p(a_2)}{\sum_{j=1,2} p(b_1|a_j)p(a_j)} = 0 \tag{14}$$

$$p(a_1|b_2) = \frac{p(b_2|a_1)p(a_1)}{\sum_{j=1,2} p(b_2|a_j)p(a_j)} = \frac{p(b_2|a_1)p(a_1)}{p(b_2|a_1)p(a_1) + p(b_2|a_2)p(a_2)} = \frac{1}{3} \tag{15}$$

$$p(a_2|b_2) = \frac{p(b_2|a_2)p(a_2)}{\sum_{j=1,2} p(b_2|a_j)p(a_j)} = \frac{p(b_2|a_2)p(a_2)}{p(b_2|a_1)p(a_1) + p(b_2|a_2)p(a_2)} = \frac{2}{3} \tag{16}$$

が得られるので, これらを (12) に代入して

$$H(\mathcal{A}|\mathcal{B}) \quad = \quad -\frac{1}{3}\times\frac{3}{4}\log\frac{1}{3}-\frac{2}{3}\times\frac{3}{4}\log\frac{2}{3}=\frac{3}{4}\log 3-\frac{1}{2} \tag{17}$$

となる. 従って, 相互情報量 $I(\mathcal{A};\mathcal{B})$ は

$$I(\mathcal{A};\mathcal{B}) \quad = \quad 1-\left\{\frac{3}{4}\log 3-\frac{1}{2}\right\}=\frac{3}{2}-\frac{3}{4}\log 3\simeq 0.31 \ (\text{ビット}) \tag{18}$$

となる.

2. ラグランジュの未定係数法を用いることにより, n 値エントロピー関数の最大値とその最大値を与える確率を求めていく. 問題文の誘導に従って行けばよい.

(1)

$$F(\{p_k\}) \quad = \quad H+\lambda\left(1-\sum_{i=1}^{n}p_i\right)=-\sum_{i=1}^{n}p_i\log p_i-\lambda\sum_{i=1}^{n}p_i+\lambda \tag{19}$$

に対して, $F(\{p_k\})$ の極値条件より

$$\frac{\partial F}{\partial p_k} \quad = \quad -p_k\cdot\frac{1}{p_k}-\log p_k-\lambda=0 \tag{20}$$

従って, この方程式の解である p_k^* は

$$p_k^* \quad = \quad 2^{-(1+\lambda)}\equiv\mathcal{P}(\lambda) \tag{21}$$

となる. なお,

$$\left.\frac{\partial^2 F}{\partial p_k\partial p_l}\right|_{p_k=p_k^*,p_l=p_l^*}=0, \quad \left.\frac{\partial^2 F}{\partial p_k^2}\right|_{p_k=p_k^*} \quad = \quad -\frac{1}{p_k^*}=-2^{(1+\lambda)}<0 \quad (\forall_{k,l\neq k}) \tag{22}$$

であるから, p_k^* は F の最大値を与える.

(2) 全ての k について, p_k^* を足しあげたものは 1 にならなければならないので

$$\sum_{k=1}^{n}p_k^* \quad = \quad 2^{-(1+\lambda)}\sum_{k=1}^{n}1=n\cdot 2^{-(1+\lambda)}=1 \tag{23}$$

が成り立つべきであり, これを p_k^*, 及び, λ に関して解けば

$$p_k^*=\frac{1}{n}, \quad \lambda=\log n-1 \tag{24}$$

が得られる. 従って, これらを n 値エントロピー関数 H に代入すれば, H の最大値は

$$H(\{p_k^*\}) \quad = \quad -\sum_{k=1}^{n}p_k^*\log p_k^*=-\frac{1}{n}\times n\log\frac{1}{n}=\log n \tag{25}$$

となる. ここで, $n=2$ とおけば教科書の 2 値エントロピー関数の場合と一致することに注意.

(別解)

この問題では n 値エントロピー関数を最大にする確率が, n 個全ての事象が等確率で現れるとき, すなわち $p_1^* = p_2^* = \cdots = p_n^* = 1/n$ で与えられることは直観的にすぐわかることですから, この解 $p_1^* = p_2^* = \cdots = p_n^* = 1/n$ が実際に n 値エントロピー関数 $H(\{p_k\})$ を最大にすることを「示す」ことによって, ラグランジュの未定係数法を用いることなく解答を作ることもできます. ただし, ここでも確率の規格化条件：

$$\sum_{j=1}^{n} p_j \quad = \quad 1 \tag{26}$$

あるいは同じことですが

$$p_n \quad = \quad 1 - \sum_{k=1}^{n-1} p_k \tag{27}$$

が成り立っていることを忘れないようにしなければなりません.

さて, ここで具体的に示すべきことは次の 2 点です.

- n 値エントロピー関数 H の p_k $(k = 1, \cdots, n)$ に関する 1 階微分係数が $p_k^* = 1/n$ でゼロとなる.
- H の p_k に関する 2 階微分係数の作る $n \times n$ の行列式の $p_k = p_k^*$ $(k = 1, \cdots, n)$ での値：

$$I^{(n)} \quad = \quad \begin{vmatrix} I_{11} & I_{12} & \cdots & I_{1n} \\ I_{21} & I_{22} & \cdots & I_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ I_{n1} & I_{n2} & \cdots & I_{nn} \end{vmatrix} \tag{28}$$

 が行列式のサイズを 1 つずつ増加させるにつれ, その符号を変えて行く. 言い方をかえれば, n 以下の任意の自然数 m に対し, $I^{(m)}$ は $(-1)^m$ のような因子を持つ. ここで各要素 I_{kl} は $I_{kl} = (\partial^2 H/\partial p_k \partial p_l)$ の $p_k^* = p_l^* = 1/n$ での値である.

まずは 1 番目の条件ですが, 条件 (27) があるので

$$\frac{\partial p_n}{\partial p_k} \quad = \quad -1 \tag{29}$$

が成り立つことに注意し, n 値エントロピー関数を

$$\begin{aligned} H \quad &= \quad -\sum_{k=1}^{n} p_k \log p_k \\ &= \quad -\sum_{j=1}^{n-1} p_j \log p_j - p_n \log p_n \end{aligned} \tag{30}$$

と書き直しておけば

$$\begin{aligned} \frac{\partial H}{\partial p_k} \quad &= \quad -1 - \log p_k - \frac{\partial}{\partial p_n}\{p_n \log p_n\} \left(\frac{\partial p_n}{\partial p_k}\right) \\ &= \quad \log\left(\frac{p_n}{p_k}\right) \end{aligned} \tag{31}$$

つまり, $k=1,\cdots,n-1$ に対し, $p_k=p_n$, すなわち, $p_1^*=p_2^*=\cdots=p_n^*=1/n$ で H が極値をとることが示せます.

次は 2 番目の条件, つまり, $p_1^*=p_2^*=\cdots=p_n^*=1/n$ で H が最大値をとることの証明ですが, 実際それをここで示す前に, この条件の成立は, 本問題でラグランジュの未定係数法を用いる際に導入した関数 F の場合には自明であったことに注意しておきましょう. なぜならば, 関数 F には確率の規格化条件 (27) が既に取り込んであるので, F の中で変数 $p_1,\cdots,p_n$ は全て独立となり, F の 2 階微分からなる行列の非対角要素 $(\partial^2 F/\partial p_k \partial p_l)$ $(k\neq l)$ は全てゼロであり, m 次の行列式は m 個の対角要素を全て掛け合わせればよいだけで, 確かに $\{-2\lambda^{1+\lambda}\}^m \sim (-1)^m$ の因子を持つことが直ちにわかるからです.
一方, ここで考える n 値エントロピー関数 H の場合, これはさほど自明ではありません. というのも

$$
\begin{aligned}
I_{kl} &= \frac{\partial^2 H}{\partial p_k \partial p_l} \\
&= -\frac{1}{p_k}\left(\frac{\partial p_k}{\partial p_l}\right)+\frac{1}{p_n}\left(\frac{p_n}{\partial p_l}\right) \\
&= \begin{cases} -2n & (k=l) \\ -n & (k\neq l) \end{cases}
\end{aligned}
\tag{32}
$$

となり, 非対角要素が残るからです. 従って, 前出 2 番目の条件を示すために我々が考えるべき行列式は

$$
I^{(n)} = \begin{vmatrix} -2n & -n & \cdots & -n \\ -n & -2n & \cdots & -n \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ -n & \cdots & \cdots & -2n \end{vmatrix} \tag{33}
$$

であり, $m\leq n$ なる m に対し, $I^{(m)}$ を計算し, これが $(-1)^m$ なる因子を持つことを示すことがここでの目標となります.
これは一見すると厄介そうですが, 行列式の基本変形と余因子展開を用いて次のように変形すれば簡単な漸化式が得られます.

$$
I^{(m)} = \begin{vmatrix} -n & -n & \cdots & -n & -n \\ n & -2n & -n & \cdots & -n \\ 0 & -n & \cdots & \cdots & -n \\ \cdots & \cdots & \cdots & \cdots & -n \\ \cdots & \cdots & \cdots & \cdots & -n \\ 0 & -n & \cdots & \cdots & -2n \end{vmatrix} = (-n)I^{(m-1)}+(-n)K^{(m-1)} \tag{34}
$$

ここで行列式 $K^{(m-1)}$ は

$$
K^{(m-1)} = \begin{vmatrix} -n & -n & \cdots & -n \\ -n & -2n & \cdots & -n \\ \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ -n & -n & \cdots & -2n \end{vmatrix} = (-n)K^{(m-2)} \tag{35}
$$

で定義されます. 従って我々は次の連立漸化式：

$$
\begin{aligned}
I^{(m)} &= (-n)I^{(m-1)} + (-n)K^{(m-1)} &(36)\\
K^{(m-1)} &= (-n)K^{(m-2)} &(37)
\end{aligned}
$$

を解けばよいことになったわけです. (37) は直ちに

$$K^{(m-1)} = (-n)^2 K^{(m-3)} = \cdots = (-n)^{(m-2)} K^{(1)} \tag{38}$$

となり, これを用いれば (36) も同様にして, 逐次的に

$$
\begin{aligned}
I^{(m)} &= (-n)I^{(m-1)} + (-n)^{m-1}K^{(1)}\\
&= (-n)\{-nI^{(m-2)} + (-n)^{m-2}K^{(1)}\} + (-n)^{m-1}K^{(1)}\\
&= (-n)^2 I^{(m-2)} + 2(-n)^{m-1}K^{(1)}\\
&= \cdots = (-n)^{(m-1)}\{I^{(1)} + (m-1)K^{(1)}\} &(39)
\end{aligned}
$$

と変形できるので, $I^{(1)} = K^{(1)} = -2n$ に注意すれば

$$I^{(m)} = m\,2^{m-1} n^m \,\underline{(-1)^m} \tag{40}$$

が得られ, 題意が示せました.

これにたどり着くまでに面倒な手続きが必要でしたが, もしもラグランジュの未定係数法というものを知らなければ, このような手順を踏まなければならないということであり, このことは逆に, ここで **(別解)** を例として取り上げることにより, ラグランジュの未定係数法がとても有効な方法であることが実感できたのではないかと思います.

3. $\boxed{\textbf{演習問題 1}}$ で扱った通信路の相互情報量を求めてみる.

(1) エントロピー, 結合エントロピー, 条件付きエントロピーをそれぞれ求めていこう.

$$H(\mathcal{A}) = -\sum_{j=0,1} p_{\mathcal{A}}(j)\log p_{\mathcal{A}}(j) = -\frac{1}{2}\log\frac{1}{2} - \frac{1}{2}\log\frac{1}{2} = 1 \tag{41}$$

出力値の確率が

$$
\begin{aligned}
p_{\mathcal{B}}(0) &= \sum_{k=0,1} p_{\mathcal{B}|\mathcal{A}}(0|k)p_{\mathcal{A}}(k) = p_{\mathcal{B}|\mathcal{A}}(0|0)p_{\mathcal{A}}(0) + p_{\mathcal{B}|\mathcal{A}}(0|1)p_{\mathcal{A}}(1) = \frac{1}{2}(1-q) &(42)\\
p_{\mathcal{B}}(1) &= \sum_{k=0,1} p_{\mathcal{B}|\mathcal{A}}(1|k)p_{\mathcal{A}}(k) = p_{\mathcal{B}|\mathcal{A}}(1|0)p_{\mathcal{A}}(0) + p_{\mathcal{B}|\mathcal{A}}(1|1)p_{\mathcal{A}}(1) = \frac{1}{2}(1-q) &(43)\\
p_{\mathcal{B}}(\mathsf{x}) &= \sum_{k=0,1} p_{\mathcal{B}|\mathcal{A}}(\mathsf{x}|k)p_{\mathcal{A}}(k) = p_{\mathcal{B}|\mathcal{A}}(\mathsf{x}|0)p_{\mathcal{A}}(0) + p_{\mathcal{B}|\mathcal{A}}(\mathsf{x}|1)p_{\mathcal{A}}(1) = q &(44)
\end{aligned}
$$

であることから

$$H(\mathcal{B}) = -\sum_{j=0,1,\mathsf{x}} p_{\mathcal{B}}(j)\log p_{\mathcal{B}}(j) = 1 - q - (1-q)\log(1-q) - q\log q \tag{45}$$

が得られる. また, 結合エントロピーは

$$
\begin{aligned}
H(\mathcal{A},\mathcal{B}) &= \sum_i \sum_j p_{\mathcal{A},\mathcal{B}}(i,j) \log p_{\mathcal{A},\mathcal{B}}(i,j) \\
&= -\sum_i \sum_j p_{\mathcal{B}|\mathcal{A}}(i|j) p_{\mathcal{A}}(j) \log p_{\mathcal{B}|\mathcal{A}}(i|j) p_{\mathcal{A}}(j) \\
&= -p_{\mathcal{B}|\mathcal{A}}(0|0) p_{\mathcal{A}}(0) \log p_{\mathcal{B}|\mathcal{A}}(0|0) p_{\mathcal{A}}(0) - p_{\mathcal{B}|\mathcal{A}}(1|0) p_{\mathcal{A}}(0) \log p_{\mathcal{B}|\mathcal{A}}(1|0) p_{\mathcal{A}}(0) \\
&- -p_{\mathcal{A}|\mathcal{B}}(\times|0) p_{\mathcal{A}}(0) \log p_{\mathcal{B}|\mathcal{A}}(\times|0) p_{\mathcal{A}}(0) \\
&- p_{\mathcal{B}|\mathcal{A}}(0|1) p_{\mathcal{A}}(1) \log p_{\mathcal{B}|\mathcal{A}}(0|1) p_{\mathcal{A}}(1) - p_{\mathcal{B}|\mathcal{A}}(1|1) p_{\mathcal{A}}(1) \log p_{\mathcal{B}|\mathcal{A}}(1|1) p_{\mathcal{A}}(1) \\
&- p_{\mathcal{B}|\mathcal{A}}(\times|1) p_{\mathcal{A}}(1) \log p_{\mathcal{B}|\mathcal{A}}(\times|1) p_{\mathcal{A}}(1) \\
&= 1 - (1-p-q)\log(1-p-q) - p\log p - q \log q
\end{aligned} \tag{46}
$$

となる. 条件付きエントロピーは

$$
\begin{aligned}
H(\mathcal{B}|\mathcal{A}) &= -\sum_i \sum_j p_{\mathcal{B}|\mathcal{A}}(i|j) p_{\mathcal{A}}(j) \log p_{\mathcal{B}|\mathcal{A}}(i|j) \\
&= -p_{\mathcal{B}|\mathcal{A}}(0|0) p_{\mathcal{A}}(0) \log p_{\mathcal{B}|\mathcal{A}}(0|0) - p_{\mathcal{B}|\mathcal{A}}(1|0) p_{\mathcal{A}}(0) \log p_{\mathcal{B}|\mathcal{A}}(1|0) \\
&- p_{\mathcal{A}|\mathcal{B}}(\times|0) p_{\mathcal{A}}(0) \log p_{\mathcal{B}|\mathcal{A}}(\times|0) \\
&- p_{\mathcal{B}|\mathcal{A}}(0|1) p_{\mathcal{A}}(1) \log p_{\mathcal{B}|\mathcal{A}}(0|1) - p_{\mathcal{B}|\mathcal{A}}(1|1) p_{\mathcal{A}}(1) \log p_{\mathcal{B}|\mathcal{A}}(1|1) \\
&- p_{\mathcal{B}|\mathcal{A}}(\times|1) p_{\mathcal{A}}(1) \log p_{\mathcal{B}|\mathcal{A}}(\times|1) \\
&= -(1-p-q)\log(1-p-q) - p\log p - q \log q
\end{aligned} \tag{47}
$$

である. 従って, これらの結果から

$$
H(\mathcal{A}) + H(\mathcal{B}|\mathcal{A}) = 1 - (1-p-q)\log(1-p-q) - p\log p - q\log q = H(\mathcal{A},\mathcal{B}) \tag{48}
$$

が成り立っていることが確かめられる.
一方, 条件付きエントロピー $H(\mathcal{A}|\mathcal{B})$ は

$$
\begin{aligned}
H(\mathcal{A}|\mathcal{B}) &= -\sum_i \sum_j p_{\mathcal{B}|\mathcal{A}}(i|j) p_{\mathcal{A}}(j) \log p_{\mathcal{A}|\mathcal{B}}(j|i) \\
&= -p_{\mathcal{B}|\mathcal{A}}(0|0) p_{\mathcal{A}}(0) \log p_{\mathcal{A}|\mathcal{B}}(0|0) - p_{\mathcal{B}|\mathcal{A}}(1|0) p_{\mathcal{A}}(0) \log p_{\mathcal{A}|\mathcal{B}}(1|0) \\
&- p_{\mathcal{A}|\mathcal{B}}(\times|0) p_{\mathcal{A}}(0) \log p_{\mathcal{A}|\mathcal{B}}(\times|0) \\
&- p_{\mathcal{B}|\mathcal{A}}(0|1) p_{\mathcal{A}}(1) \log p_{\mathcal{A}|\mathcal{B}}(0|1) - p_{\mathcal{B}|\mathcal{A}}(1|1) p_{\mathcal{A}}(1) \log p_{\mathcal{A}|\mathcal{B}}(1|1) \\
&- p_{\mathcal{B}|\mathcal{A}}(\times|1) p_{\mathcal{A}}(1) \log p_{\mathcal{A}|\mathcal{B}}(\times|1)
\end{aligned} \tag{49}
$$

となるが,

$$
p_{\mathcal{A}|\mathcal{B}}(0|0) = \frac{p_{\mathcal{B}|\mathcal{A}}(0|0) p_{\mathcal{A}}(0)}{p_{\mathcal{B}|\mathcal{A}}(0|0) p_{\mathcal{A}}(0) + p_{\mathcal{B}|\mathcal{A}}(0|1) p_{\mathcal{A}}(1)} = \frac{1-p-q}{1-q} \tag{50}
$$

$$
p_{\mathcal{A}|\mathcal{B}}(0|1) = \frac{p_{\mathcal{B}|\mathcal{A}}(1|0) p_{\mathcal{A}}(0)}{p_{\mathcal{B}|\mathcal{A}}(1|0) p_{\mathcal{A}}(0) + p_{\mathcal{B}|\mathcal{A}}(1|1) p_{\mathcal{A}}(1)} = \frac{p}{1-q} \tag{51}
$$

$$
p_{\mathcal{A}|\mathcal{B}}(0|\times) = \frac{p_{\mathcal{B}|\mathcal{A}}(\times|0) p_{\mathcal{A}}(0)}{p_{\mathcal{B}|\mathcal{A}}(\times|0) p_{\mathcal{A}}(0) + p_{\mathcal{B}|\mathcal{A}}(\times|1) p_{\mathcal{A}}(1)} = \frac{1}{2} \tag{52}
$$

$$
p_{\mathcal{A}|\mathcal{B}}(1|0) = \frac{p_{\mathcal{B}|\mathcal{A}}(0|1) p_{\mathcal{A}}(1)}{p_{\mathcal{B}|\mathcal{A}}(0|0) p_{\mathcal{A}}(0) + p_{\mathcal{B}|\mathcal{A}}(0|1) p_{\mathcal{A}}(1)} = \frac{p}{1-q} \tag{53}
$$

$$
p_{\mathcal{A}|\mathcal{B}}(1|1) = \frac{p_{\mathcal{B}|\mathcal{A}}(1|1) p_{\mathcal{A}}(1)}{p_{\mathcal{B}|\mathcal{A}}(1|0) p_{\mathcal{A}}(0) + p_{\mathcal{B}|\mathcal{A}}(1|1) p_{\mathcal{A}}(1)} = \frac{1-p-q}{1-q} \tag{54}
$$

$$p_{\mathcal{A}|\mathcal{B}}(1|\mathsf{x}) = \frac{p_{\mathcal{B}|\mathcal{A}}(\mathsf{x}|1)p_{\mathcal{A}}(1)}{p_{\mathcal{B}|\mathcal{A}}(\mathsf{x}|0)p_{\mathcal{A}}(0)+p_{\mathcal{B}|\mathcal{A}}(\mathsf{x}|1)p_{\mathcal{A}}(1)} = \frac{1}{2} \tag{55}$$

であることに注意すると

$$H(\mathcal{A}|\mathcal{B}) = q-(1-p-q)\log(1-p-q)-p\log p+(1-q)\log(1-q) \tag{56}$$

が得られ, 関係式

$$H(\mathcal{B})+H(\mathcal{A}|\mathcal{B}) = 1-(1-p-q)\log(1-p-q)-p\log p-q\log q = H(\mathcal{A},\mathcal{B}) \tag{57}$$

が成り立っていることが確かめられる.

(3) (2) で得られた結果から相互情報量 $I(\mathcal{A};\mathcal{B})$ は

$$\begin{aligned} I(\mathcal{A};\mathcal{B}) &= H(\mathcal{A})-H(\mathcal{A}|\mathcal{B}) \\ &= 1-q+(1-p-q)\log(1-p-q)+p\log p-(1-q)\log(1-q) \end{aligned} \tag{58}$$

となるが, $q=0$ とおけば (これは教科書の 2 元対称通信路の場合)

$$I(\mathcal{A};\mathcal{B}) = 1+(1-p)\log(1-p)+p\log p \tag{59}$$

となり, p に関する極値条件 $(\partial I/\partial p)=0$ から直ちに

$$p = \frac{1}{2} \tag{60}$$

が得られるが, $(\partial^2 I/\partial p^2)|_{p=1/2}=4>0$ であるから, この $p=1/2$ で相互情報量は最小となる. この通信路の相互情報量は「出力値を観測することによって減少する入力値に関する『あいまいさ』の量」であるから, 2 元対称通信路の場合, 確率 1/2 で入力信号が反転してしまえば, 出力値を知ることにより入力値に関してもたらされる情報量は最小 (ゼロ) となる. これは直観と照らし合わせて, もっともらしい. なお, この相互情報量は $p=0,1$ のとき, すなわち, 「全くノイズが無い」「全て反転する」場合につき最大値をとることも容易に見て取れるであろう. 図 1 にここでの相互情報量 I

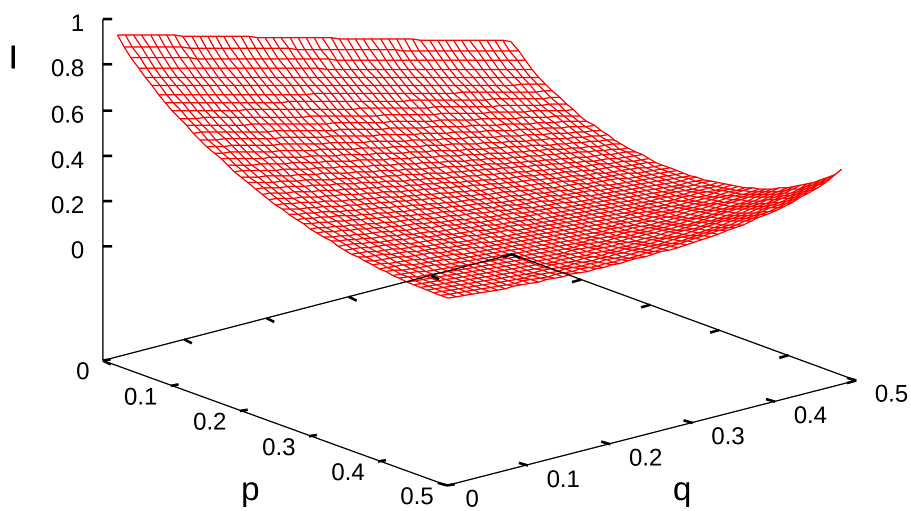

図 1: この問題で扱う通信路の相互情報量 I. p,q はぞれぞれ $[0,0.5]$ の間でプロットしてある. $q=0$ のとき, I は $p=1/2$ で最小値をとる (I は $p=1/2$ の軸に関して対称であることに注意).

を p,q の関数としてプロットしておこう.

演習問題 3

情報源 X の**エントロピーレート**：

$$\mathcal{H}(X) = \lim_{n\to\infty} \frac{1}{n} H(X_1, X_2, \cdots, X_n) \tag{61}$$

に関して以下の問いに答えよ.

(1) 記号 $0, 1$ をそれぞれ $p, 1-p$ で生成する**定常的無記憶情報源**のエントロピーレートを求めよ.

(2) 定常分布 $p(x)$ が

$$p(x) = p\,\delta_{x,1} + (1-p)\,\delta_{x,0} \tag{62}$$

状態遷移確率 $p(x^{'}|x)$ が

$$p(x^{'}|x) = q + (1-2q)\,\delta_{x,x'} \tag{63}$$

で与えられる**定常的単純マルコフ情報源**を考えよう ($\delta_{a,b}$ はクロネッカ・デルタである). このとき

- この状態遷移確率 $p(x^{'}|x)$ を表す状態遷移図.
- この情報源のエントロピーレート.

のそれぞれを求めよ.

次回からいよいよ「情報源符号化 (圧縮)」に入ります.
http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/InfoTheory/InfoTheory04_report.html でレポートの提出状況が確認できます.