講義「情報理論」

第5回　情報源のモデル（後半）

情報理工学専攻 情報知識ネットワーク研究室
喜田拓也

2017/7/25

離散的M元情報源（おさらい）

M 個の元からなる記号の有限集合 $A = \{a_1, a_2, \ldots, a_M\}$ を考える.
これを情報源アルファベットとよび，各記号を情報源記号という

情報源は，時点0より毎時点（時点は整数値）において，情報源アルファベットA上の記号をある確率に従って1個ずつ出力する.
このような情報源を，離散的M元情報源という

離散的46元情報源
$A = \{$あ, い, う, $\cdots$, わ, を, ん$\}$

記号を1個ずつ出力

$\Rightarrow$

0	1	2	3	4	
か	き	ふ	ら	い	・・・・・・

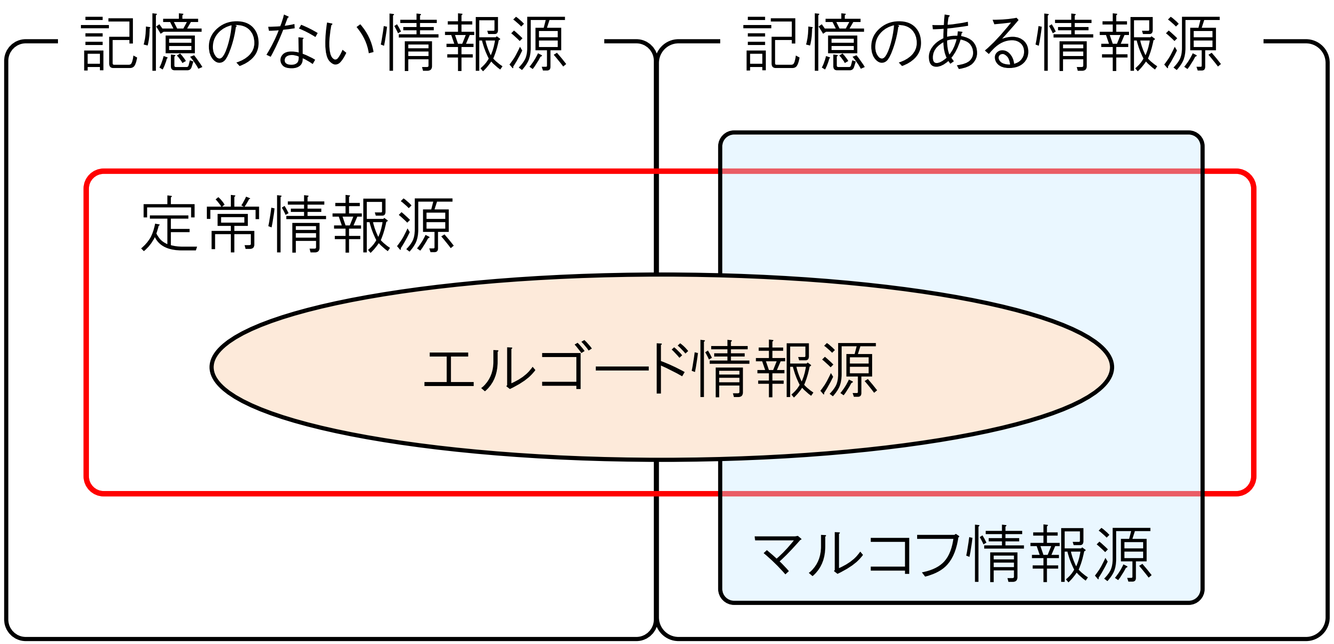
情報源系列

情報源系列の統計的性質は結合確率分布で定まる！

情報源のモデル（おさらい）

どんな大きいNについても, $X_0, X_1, \cdots, X_{N-1}$の結合確率分布を与えることができれば, この情報源の統計的性質は完全に記述できる. しかし, 一般には困難.

議論を進めるためには, 情報源に何らかの扱いやすい性質を仮定する必要がある.

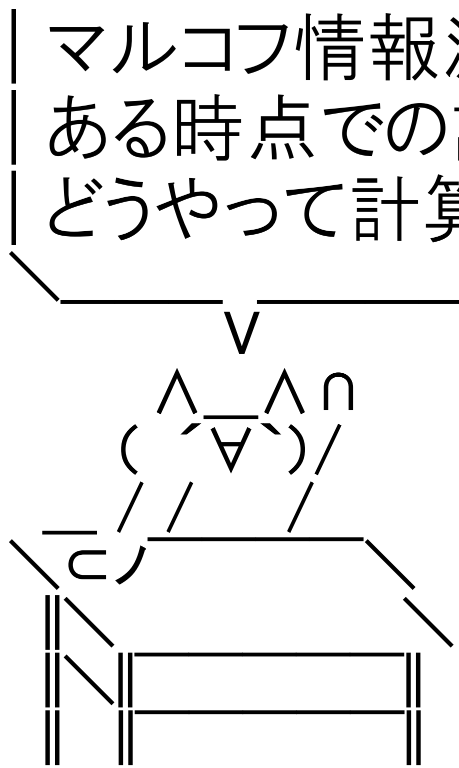


今日の内容

3.4 マルコフ情報源の確率分布

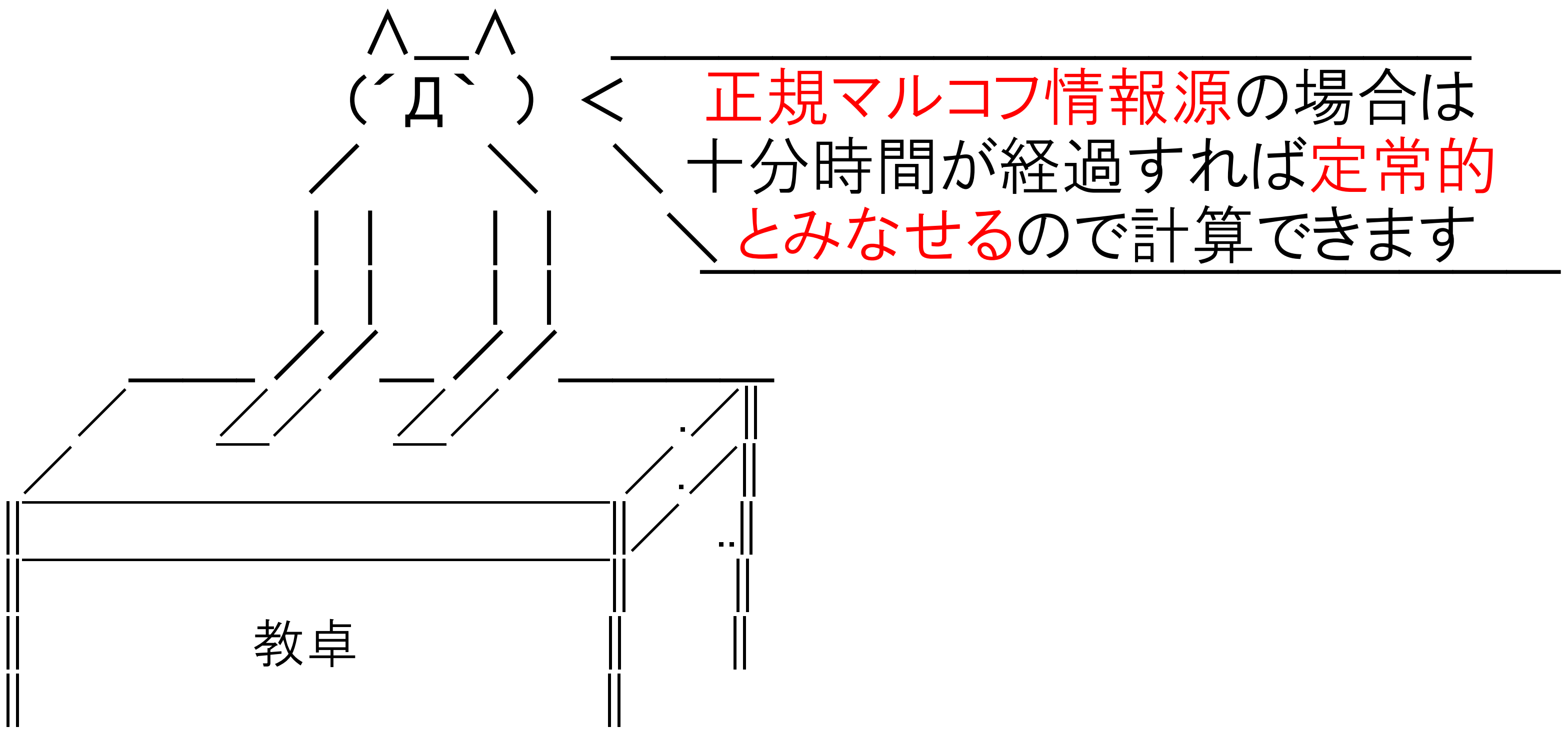
3.5 情報源のエントロピー

素朴な疑問



どこにいるか分かんないじゃ…

なんと！



遷移確率行列

N個の状態$s_0, s_1, \ldots, s_{N-1}$を持つ正規マルコフ情報源を考える.

状態遷移の仕方は，状態s_iにあるとき，次の時点で状態s_jに遷移する確率 $p_{i,j} = P\left(s_j \middle| s_i\right)$により決まる. これを遷移確率という.

遷移確率$p_{i,j}$を(i,j)要素とする$N \times N$行列を遷移確率行列と呼ぶ.

遷移確率行列 $\Pi = \begin{pmatrix} p_{0,0} & \cdots & p_{0,N-1} \\ \vdots & \ddots & \vdots \\ p_{N-1,0} & \cdots & p_{N-1,N-1} \end{pmatrix}$

行 $= i$ 側
列 $= j$ 側

行ごとに
総和が1

例題3.5

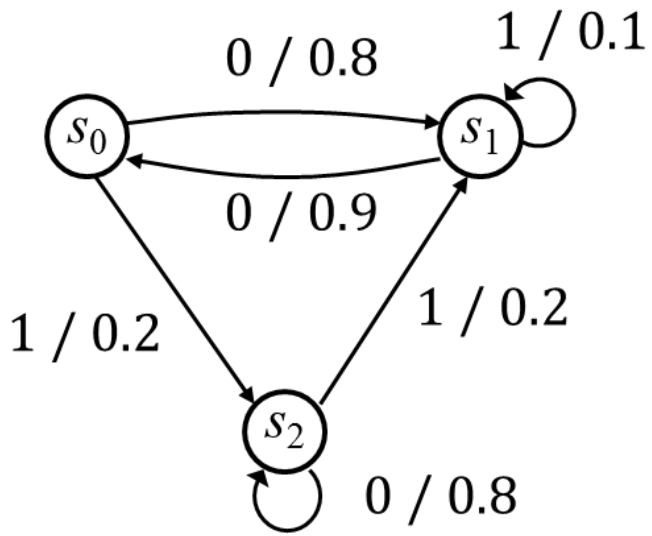


図3.6

図3.6の状態図に従うマルコフ情報源の遷移確率行列Πを求めよ.

$$\Pi = \begin{array}{c} \begin{pmatrix} 0 & 0.8 & 0.2 \\ 0.9 & 0.1 & 0 \\ 0 & 0.2 & 0.8 \end{pmatrix} \\ \underbrace{s_0 \qquad s_1 \qquad s_2}_{\text{へ}} \end{array} \left.\begin{array}{l} \leftarrow s_0 \\ \leftarrow s_1 \\ \leftarrow s_2 \end{array}\right\} \text{から}$$

行ごとに
総和が1

Try 練習問題3.4

遷移確率行列による t 時点後の遷移確率

$$\Pi = \begin{pmatrix} p_{0,0} & \cdots & p_{0,N-1} \\ \vdots & \ddots & \vdots \\ p_{N-1,0} & \cdots & p_{N-1,N-1} \end{pmatrix}$$

状態数は N個！

状態s_i から出発し, t 時点後に s_j に到達する確率を$p_{i,j}^{(t)}$とする.

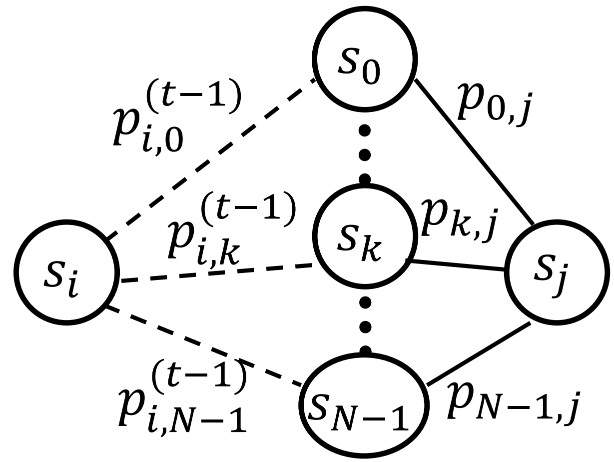
$p_{i,j}^{(1)} = p_{i,j}$は明らか.

状態s_iから $t-1$ 時点後に状態s_kを通って

状態s_jに行く確率は $p_{i,k}^{(t-1)} \cdot p_{k,j}$ なので,

$$p_{i,j}^{(t)} = \sum_{k=0}^{N-1} p_{i,k}^{(t-1)} \cdot p_{k,j}$$

により計算できる.



行列の掛算

$$\underbrace{\Pi \cdot \Pi \cdots \Pi}_{t\ 回}$$

この式から $p_{i,j}^{(t)} = (\Pi^t の(i,j)要素)$ となる.

例題3.6

図3.6の状態遷移図で表される2元マルコフ情報源を考える.
時点0で状態s_0からスタートしたとして，時点2のときに状態s_2にいる確率はいくらになるか求めよ.

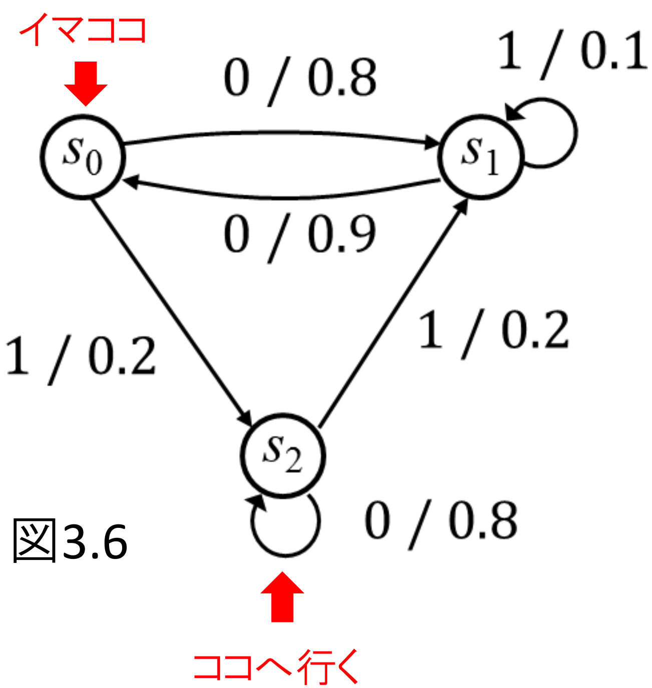


図3.6

状態遷移行列は先の例題より,

$$\Pi = \begin{pmatrix} 0 & 0.8 & 0.2 \\ 0.9 & 0.1 & 0 \\ 0 & 0.2 & 0.8 \end{pmatrix}.$$

このΠ から，$p_{0,2}^{(2)}$ を求めたい．すなわち，

Π^2の(0,2)の要素を求めればよい．

$$\Pi^2 = \begin{pmatrix} 0.72 & 0.12 & 0.16 \\ 0.09 & 0.73 & 0.18 \\ 0.18 & 0.18 & 0.64 \end{pmatrix}$$

より，$p_{0,2}^{(2)} = 0.16$ である． Try 練習問題3.5

正規マルコフ情報源の極限分布

正規マルコフ情報源の定義より, $t > t_0$ となる任意の t について

$$p_{i,j}^{(t)} > 0 \quad (i, j = 0, 1, \cdots, N-1)$$

が成り立つようなある正定数 t_0 が存在する.

十分時間が経てば,
どの状態にも
遷移し得るから

正規マルコフ情報源では, $t \to \infty$ とするとき,

$p_{i,j}^{(t)}$ は i には無関係な値に収束する[証明は省略]. すなわち,

$$\lim_{t\to\infty} p_{i,j}^{(t)} = u_j \quad (j = 0, 1, \cdots, N-1)$$

j だけに
依存する値

となる u_j が存在する.

遷移行列を用いて上式を書き直すと

$$\lim_{t\to\infty} \Pi^t = U$$

となる. ここで, U はすべての行が $\boldsymbol{u} = (u_0, u_1, \cdots, u_{N-1})$

となる $N \times N$ 行列である.

正規マルコフ情報源の極限分布(つづき)

時点 t において状態 s_j にいる確率$P\left(s^{(t)} = s_j\right)$ を $w_j^{(t)}$で表し,

$$\boldsymbol{w}_t = (w_0^{(t)}, w_1^{(t)}, \cdots, w_{N-1}^{(t)})$$

という N次元ベクトルを定義する.

状態確率分布ベクトル
または状態分布

時点$t-1$で状態s_iにいる確率が$w_i^{(t-1)}$で,状態s_iにいるときに次の時点で状態s_jへと遷移する確率が$p_{i,j}$であるから,

$$w_j^{(t)} = \sum_{i=0}^{N-1} w_i^{(t-1)} p_{i,j}$$

$$\therefore \boldsymbol{w}_t = \boldsymbol{w}_{t-1}\, \Pi\ .$$

これを繰り返せば $\boldsymbol{w}_t = \boldsymbol{w}_0\, \Pi^t$ を得る.ここで,$\boldsymbol{w}_0$は時点0における状態分布(初期分布)である.

$\boldsymbol{w}_t$ の $t \to \infty$ とした極限を極限分布と呼ぶ.

極限分布は一般に
存在するとは限らない

正規マルコフ情報源では,次式が成り立つ.

$$\lim_{t\to\infty} \boldsymbol{w}_t = \boldsymbol{w}_0 \lim_{t\to\infty} \Pi^t = \boldsymbol{w}_0 U = \boldsymbol{u}$$

初期分布によらない

正規マルコフ情報源の定常分布

十分時間が経過すれば，初期分布がどうであれ，状態分布は定常的な確率分布（定常分布）に落ち着く.

正規マルコフ情報源が落ち着く定常分布を

$$\boldsymbol{w} = (w_0, w_1, \cdots, w_{N-1})$$

とする. w_iは確率なので，当然ながら

大事な式！

ある時点の状態分布が定常的で $\boldsymbol{w}$ であるとすれば，次の時点の状態分布も $\boldsymbol{w}$ でなければならないので，$\boldsymbol{w}$ は

大事な式！

を満たさなければならない.

正規マルコフ情報源の遷移確率行列 Π に対しては，この式を満たす $\boldsymbol{w}$ が唯一存在し，極限分布と一致する.

例題3.7

図3.6で表現される情報源S の定常分布 $\boldsymbol{w} = (w_0, w_1, w_2)$ を求めよ．また，情報源S が定常分布にあるときの各記号の生起確率を求めよ．

（方針） まず，先に述べた次の二つの関係式から定常分布を求める．

$$\begin{cases} w_0 + w_1 + \ \cdots + w_{N-1} = 1, \\ \boldsymbol{w}\Pi = \boldsymbol{w} \end{cases}$$

その後，各記号の生起確率を定常分布の割合に従って計算する．

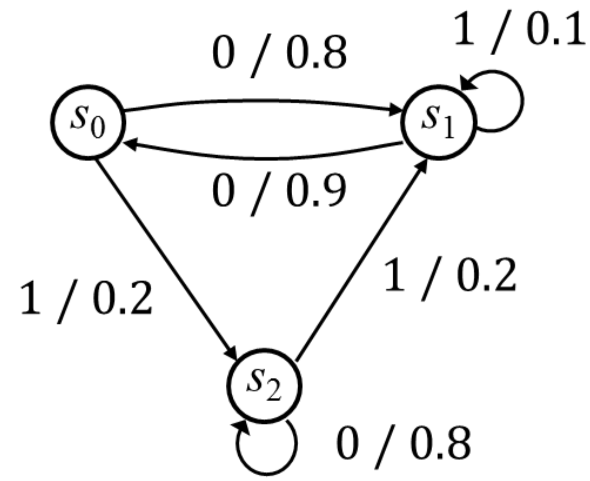


図3.6

状態遷移確率Π

$$\Pi = \begin{pmatrix} 0 & 0.8 & 0.2 \\ 0.9 & 0.1 & 0 \\ 0 & 0.2 & 0.8 \end{pmatrix}$$

例題3.7（つづき）

定常状態の満たすべき式 $\boldsymbol{w}\Pi = \boldsymbol{w}$ より,

$$\begin{aligned} 0 \quad + 0.9\,w_1 + \quad 0 \quad &= w_0, \\ 0.8\,w_0 + 0.1\,w_1 + 0.2w_2 \; &= w_1, \\ 0.2w_0 \; + \quad 0 \quad + 0.8\,w_2 &= w_2. \end{aligned}$$

状態遷移確率Π

$$\Pi = \begin{pmatrix} 0 & 0.8 & 0.2 \\ 0.9 & 0.1 & 0 \\ 0 & 0.2 & 0.8 \end{pmatrix}$$

また, $w_0 + w_1 + w_2 = 1$である. これらの連立方程式を解くと,

$$(w_0, w_1, w_2) = \left(\frac{9}{28}, \frac{10}{28}, \frac{9}{28}\right).$$

情報源S が定常分布にあるとき, 0,1が出力される確率は各々,

$$P(0) = \frac{9}{28} \times 0.8 + \frac{10}{28} \times 0.9 + \frac{9}{28} \times 0.8 = \frac{234}{280} = \frac{117}{140},$$

$$P(1) = 1 - P(0) = \frac{23}{140}$$

である.

Try 練習問題3.6

ちょっと休憩

エントロピーの定義(おさらい)

定義2.3

確率変数Xがとりうる値が$x_1, x_2, \cdot\cdot\cdot, x_M$とし，$X$がそれぞれの値をとる確率が$p_1, p_2, \ldots, p_M$(ただし，$p_1 + p_2 + \cdots + p_M = 1$)であるとき，確率変数$X$のエントロピーを

$$H(X) = -\sum_{i=1}^{M} p_i \log_2 p_i$$

ビットと定義する．

Xの平均情報量

<u>例題2.1</u>： 偏りのないコインを2回投げて表の出た枚数を確率変数Xとする．このとき，Xのエントロピー$H(X)$は何ビットか？

$$H(X) = -\frac{1}{4}\log_2\frac{1}{4} - \frac{1}{2}\times\log_2\frac{1}{2} - \frac{1}{4}\log_2\frac{1}{4}$$
$$= 2\times\frac{2}{4} + \frac{1}{2} = 1.5 \quad (\text{ビット})$$

X	0	1	2
確率	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

情報源の1次エントロピー

次のようなM元定常情報源Sを考える

情報源アルファベット$A = \{a_1, a_2, \cdots, a_M\}$

記号a_kの生起確率 p_k

このとき，

を情報源Sの１次エントロピーと呼ぶ．

記憶のない定常情報源の場合，その1次エントロピーは，情報源からの出力を1個受け取ったときに得られる平均的な情報量と考えることができる．

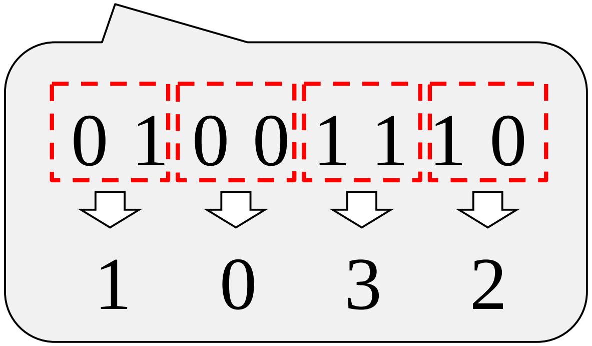
じゃあ記憶のあるときは？

情報源のn次エントロピー

長さnの系列を考え，その系列全体のエントロピーから1記号あたりのエントロピーを算出することを考える.

M元情報源のn個の出力をまとめて1個の記号とみなすと，M^n元の情報源とみなせる.

これをn次拡大情報源といい，S^nで表す.



このとき，

$$H_n(S) = \frac{H_1(S^n)}{n}$$

を情報源Sのn次エントロピーと呼ぶ.

記憶のない定常情報源の場合，そのn次エントロピーは，1次エントロピーと一致する.

情報源のエントロピー

情報源Sのn次エントロピーの極限

を情報源Sの**エントロピー**と呼ぶ.

これは，十分長い時間をかけて系列を観測したときの1記号あたりの平均情報量と考えられる.

記憶のない定常情報源の場合，
そのエントロピーは1次エントロピーと一致する.

じゃあマルコフ情報源のときは？

正規マルコフ情報源のエントロピー

（例題3.8）右図のマルコフ情報源Sのエントロピーを考える.
情報源Sの定常分布(w_0, w_1)は,

$$(w_0, w_1)\begin{pmatrix}0.9 & 0.1\\ 0.4 & 0.6\end{pmatrix} = (w_0, w_1),$$

$$w_0 + w_1 = 1$$

より, $(w_0, w_1) = (0.8, 0.2)$ と求まる.

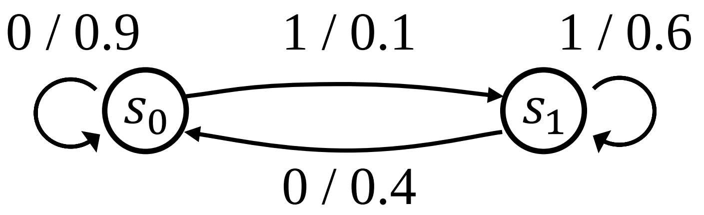


図3.10 マルコフ情報源

いま, Sが状態 s_0 にあるときだけに注目すると,
この情報源は1,0を0.1, 0.9の確率で発生する記憶のない情報源とみなせる.
その場合のエントロピーを $H_{s_0}(S)$ と書くと,

$$H_{s_0}(S) = \mathcal{H}(0.1) \fallingdotseq 0.4690 .$$

同様に, Sが状態 s_1 にあるときだけを注目すれば,

$$H_{s_1}(S) = \mathcal{H}(0.6) \fallingdotseq 0.9710 .$$

定常分布では, s_0にいる確率が$w_0 = 0.8$, s_1にいる確率が$w_1 = 0.2$だから,

$$H(S) \fallingdotseq 0.8 \times 0.4690 + 0.2 \times 0.9710 = 0.5694$$

になると考えられる.

正規マルコフ情報源のエントロピー(つづき)

正規マルコフ情報源のエントロピー

次のような正規マルコフ情報源Sを考える.

- 情報源アルファベット$\{a_1, a_2, \cdots, a_M\}$
- N個の状態 $s_0, s_1, \cdots, s_{N-1}$
- 定常分布 $(w_0, w_1, \cdots, w_{N-1})$
- 状態s_i にあるときに記号 a_kを発生する確率 $P(a_k|s_i)$

この情報源Sに対するエントロピー$H(S)$は

$$H(S) = -\sum_{i=0}^{N-1} w_i \left(\sum_{k=1}^{M} P(a_k|s_i) \log_2 P(a_k|s_i) \right)$$

正規マルコフ情報源のエントロピー(まとめ)

正規マルコフ情報源 S のエントロピー $H(S)$ の求め方の手順は次のとおり

1. 定常分布を求める
2. 各状態s_iにおいて，記憶のない情報源とみなし，その場合のエントロピー$H_{s_i}(S)$ を求める
3. 上で求めたエントロピー$H_{s_i}(S)$ を，定常分布に従った割合で合算する

Try 練習問題3.7

今日のまとめ

3.4 マルコフ情報源の確率分布

遷移確率行列

正規マルコフ情報源の極限分布

正規マルコフ情報源の定常分布

3.5 情報源のエントロピー

情報源の1次エントロピー，n次エントロピー，エントロピー

正規マルコフ情報源のエントロピーの求め方

次回：

第4章　情報源符号化とその限界