北海道大学
Hokkaido University

情報理論

第5回　情報源のモデル (2)

工学部 情報エレクトロニクス学科

情報科学研究院 情報理工学部門
大規模知識処理研究室

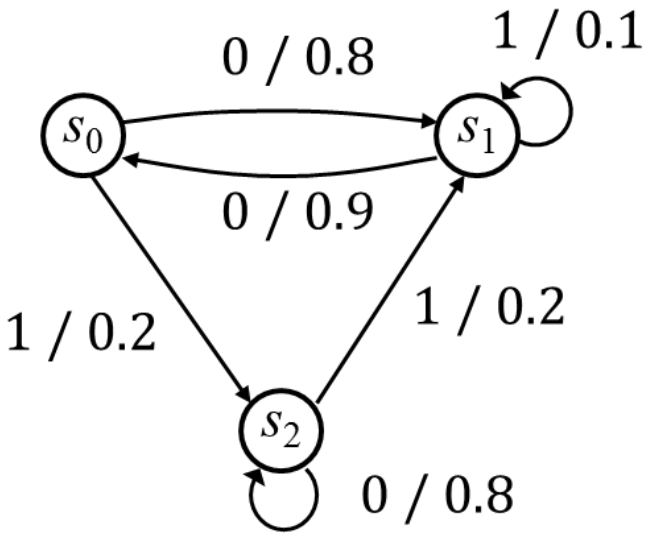
堀山 貴史

前回から今回へ: マルコフ情報源

（前回の最後）

- 正規マルコフ情報源:
  - 十分な時間が経過した後に、定常情報源として扱えて有益（次回）

今回の授業で扱う例



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 $\boldsymbol{p_{i,j} = P(s_j | s_i)}$ 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**とよぶ.

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

行 $= i$ 側
列 $= j$ 側

行ごとに
総和が1

遷移確率行列

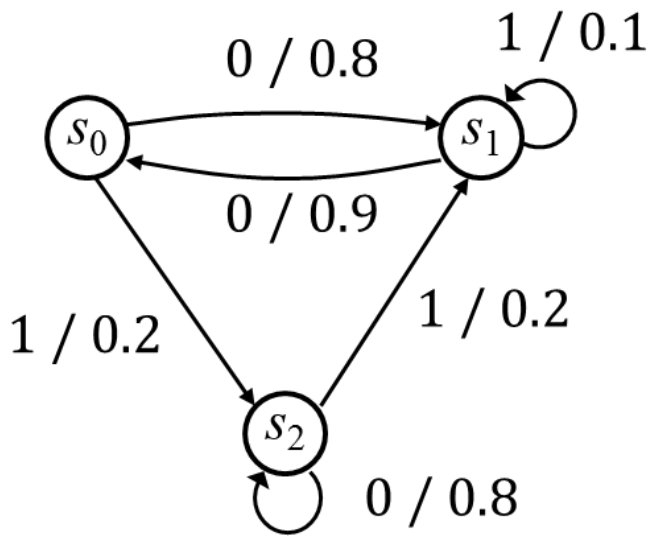


図3.6

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.

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

から

行ごとに
総和が1

遷移確率行列による t 時点後の遷移確率

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

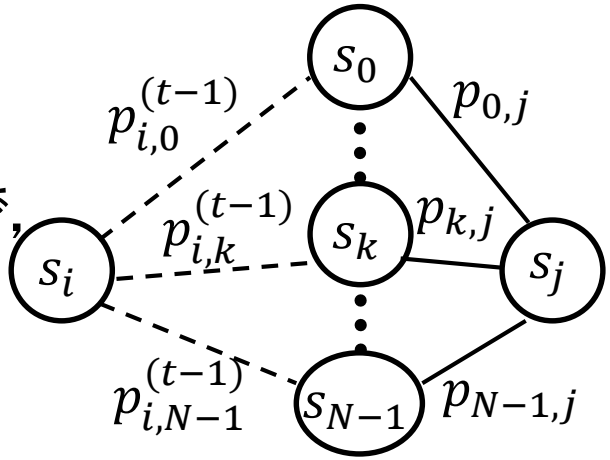
状態数は N 個

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 $t-1$ 回の遷移で状態s_kに行き

あと1回で 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

$$\underbrace{\Pi \cdot \Pi \cdots \Pi}_{t\text{ 個}}$$

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

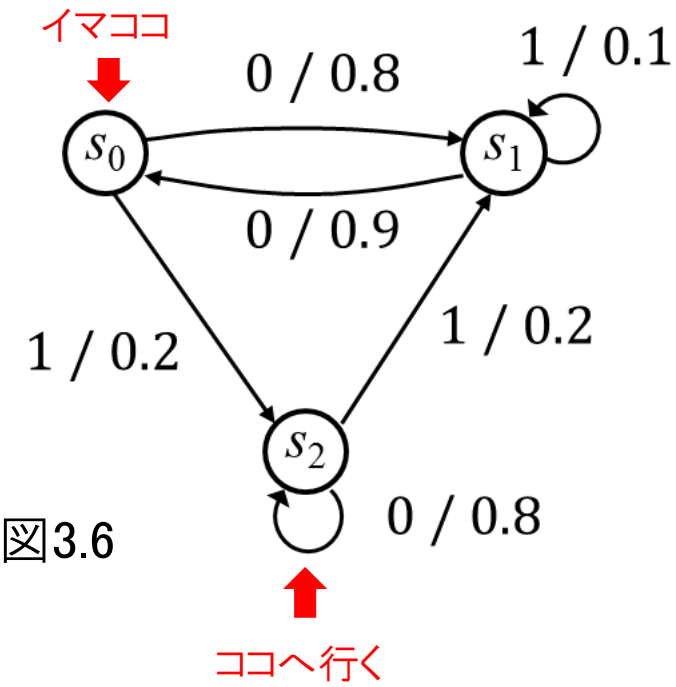


図3.6

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.

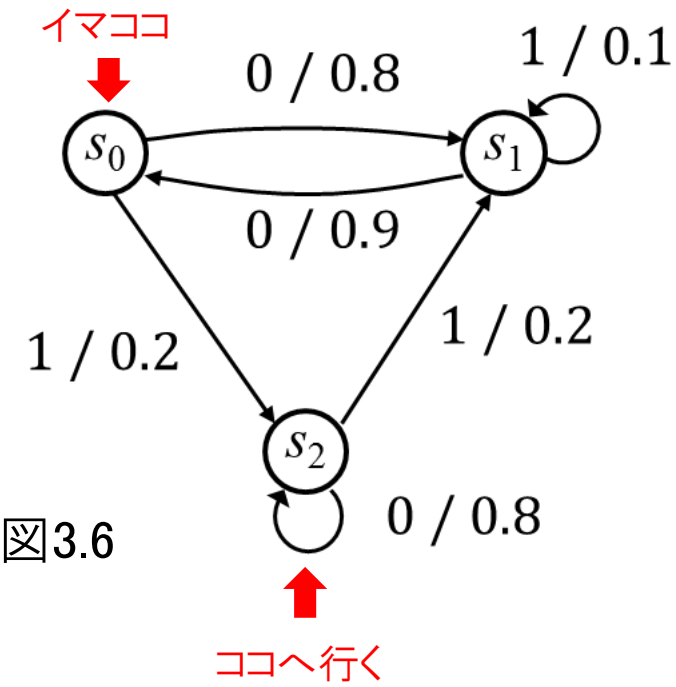


図3.6

遷移確率行列は，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.

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．

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，

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

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．

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 $w_t = w_0 \Pi^t$ 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．

正規マルコフ情報源では，次式が成り立つ．

$$\lim_{t\to\infty} \boldsymbol{w_t} = \boldsymbol{w_0} \lim_{t\to\infty} \Pi^t = \boldsymbol{w_0} U = \boldsymbol{u}$$

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

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

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．

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

$$\boldsymbol{w_0 + w_1 + \cdots + w_{N-1} = 1}\,.$$

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

$$\boldsymbol{w\Pi = w}$$

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で表現される定常2元情報源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．

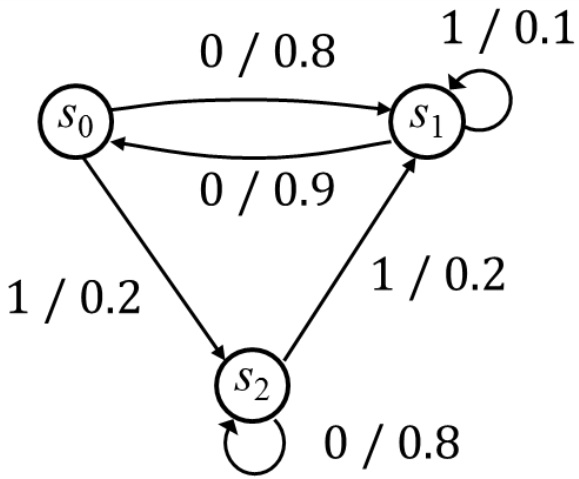


図3.6

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

図3.6で表現される定常2元情報源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．

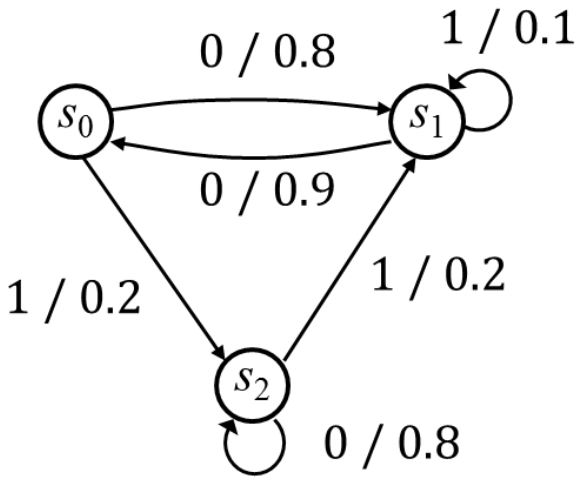


図3.6

（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．

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

状態遷移確率Π

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

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．

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，

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

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．

ちょっと休憩

情報源のエントロピー

- 前半： 正規マルコフ情報源の
  定常分布（定常的な確率分布）について、学習した
- 後半： 定常分布をもとに、**情報源のエントロピー** を求めよう

定義2.3（復習）

確率変数X がとりうる値が$x_1, x_2, \cdots, x_M$ 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

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

ビットと定義する.

Xの平均情報量

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.

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,

$$H_1(S) = -\sum_{k=1}^{M} p_k \log_2 p_k$$

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．

0 1 0 0 1 1 1 0

$n = 2$ なら、2個の出力をまとめて1 つと考える

- 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.

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

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

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

例題3.8： 図3.10の２元マルコフ情報源のエントロピー

- 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.

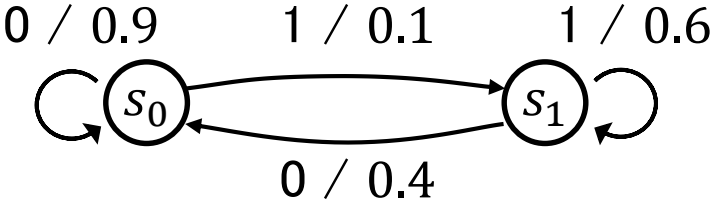


図3.10 マルコフ情報源

- 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（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

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，定常分布にしたがった割合で合算する

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