情報理論
第 4 回 確率過程

堀田 政二
工学部 情報工学科

確率過程 (stochastic process)

試行が次々と行われる行為 (時点と共に変化する確率変数) を確率過程とよぶ． e.g., **サイコロを投げ続ける行為**

- **【例 1】サイコロを投げ，t 回目に偶数の目が出たら $X_t = 0$，奇数の目が出たら $X_t = 1$ を割り当てて記録することを考える (各 X_t は互いに独立)**

  サイコロを投げ続けた時の記録の一例: $1011000100\cdots$

- **【例 2】コインを投げて表が出たら $+1$ 点，裏が出たら -1 点を獲得できるとする．t 回めのコインを投げた時に得られる得点を確率変数 X_t で表すと，t 回めのコインを投げた時までの総得点 S_t は**

$$S_t = S_{t-1} + X_t$$

  で表される．ただし，$S_0 = 0$ とする．

(2)

単純マルコフ過程 (simple Markov process)

ある時点 t における事象の生起する確率が直前の事象のみに依存するとき，単純マルコフ過程という

- **一つ前の試行の結果に応じて，記号の生起する確率が変わるサイコロ投げを考える (各 X_t は独立でない)**
  - $t-1$ **番目の記号** X_{t-1} **が** 0 **のとき，** t **回目のサイコロの目が** 5 **であれば** $X_t = 0$**，それ以外の目であれば** $X_t = 1$ **とする**
  - $t-1$ **番目の記号** X_{t-1} **が** 1 **のとき，** t **回目のサイコロの目が** 6 **であれば** $X_t = 1$**，それ以外の目であれば** $X_t = 0$ **とする．**

サイコロを投げ続けた時の記録の一例: $01010110100\cdots$

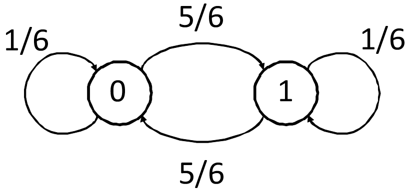
(3)

遷移確率 (transition probability) と シャノン線図 (Shannon diagram)

先ほどの単純マルコフ過程の例を条件付き確率で表すと

- $P(X_t = 0|X_{t-1} = 0) = 1/6$ (X_{t-1}=0 **で目が** 5)
- $P(X_t = 1|X_{t-1} = 0) = 5/6$ (X_{t-1}=0 **で目が** 5 **以外**)
- $P(X_t = 0|X_{t-1} = 1) = 5/6$ (X_{t-1}=1 **で目が** 6 **以外**)
- $P(X_t = 1|X_{t-1} = 1) = 1/6$ (X_{t-1}=1 **で目が** 6)

となる．通信関係の分野では，このような確率は遷移確率と呼ばれ，それを図に表したものをシャノン線図，または状態遷移図と呼ぶ．

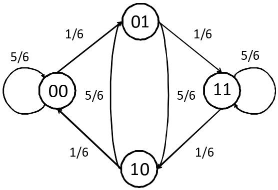


(4)

M 重マルコフ過程 (M-th order Markov process)

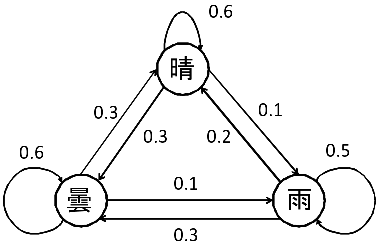
一つ前だけでなく，二つ前にも依存する確率過程も考えられる

- $P(X_t = 0 | X_{t-2} = 0, X_{t-1} = 0) = 5/6$
- $P(X_t = 1 | X_{t-2} = 0, X_{t-1} = 0) = 1/6$
- $P(X_t = 0 | X_{t-2} = 0, X_{t-1} = 1) = 5/6$
- $P(X_t = 1 | X_{t-2} = 0, X_{t-1} = 1) = 1/6$
- $P(X_t = 0 | X_{t-2} = 1, X_{t-1} = 0) = 1/6$
- $P(X_t = 1 | X_{t-2} = 1, X_{t-1} = 0) = 5/6$
- $P(X_t = 0 | X_{t-2} = 1, X_{t-1} = 1) = 1/6$
- $P(X_t = 1 | X_{t-2} = 1, X_{t-1} = 1) = 5/6$



ある時点 t における事象の生起する確率が，それより M 時点前の状態に依存するとき，この過程を M 重マルコフ過程という

天気の遷移図 (有向グラフ)



- 時点 t で状態 i であったとき，次の時点 $t+1$ に状態 j に遷移する確率は，時点 $t-1$ 以前にどの状態であったかには無関係という仮定 (マルコフ性)
- マルコフ性を持った確率過程: マルコフ過程
- 遷移確率が時点 t でも $t-1$ でもまったく同じであると仮定: 遷移確率の定常性
- 定常な遷移確率を持つマルコフ過程は斉時的 (せいじてき) という

(6)

マルコフ連鎖 (Markov chain)

マルコフ過程のうち，取りうる状態が離散的なものを特にマルコフ連鎖とよぶ

- **時点** p **の状態 (これが離散的) を** i_p $(p = 0, ..., t-1)$**，時点** t **の状態を** j **とすると，単純 (1 重) マルコフ連鎖では以下が成り立つ**

$$P(X_t = j | X_0 = i_0, ..., X_{t-1} = i_{t-1}) = P(X_t = j | X_{t-1} = i_{t-1})$$

- 2 **重マルコフ連鎖では以下が成り立つ**

$$P(X_t = j | X_0 = i_0, ..., X_{t-1} = i_{t-1}) = P(X_t = j | X_{t-2} = i_{t-2}, X_{t-1} = i_{t-1})$$

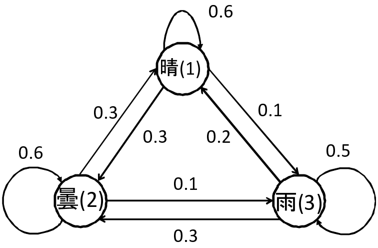
- 3 **重マルコフ連鎖では ··· 以下省略**

単純マルコフモデルにおける右辺の条件付き確率を状態遷移確率 (transition probability) **と呼び，小文字の** p_{ij} **で表す**

$$p_{ij} = P(X_t = j | X_{t-1} = i)$$

遷移確率行列 (transition probability matrix)

今日が晴のとき，m 日後晴になる確率を求めたい．そのために，遷移確率行列を導入しよう．
遷移確率行列とは，状態遷移確率を要素に持つ行列 $P = \{p_{ij}\}$ のことである．すなわち，i 行 j 列は状態 i から状態 j への状態遷移確率である．

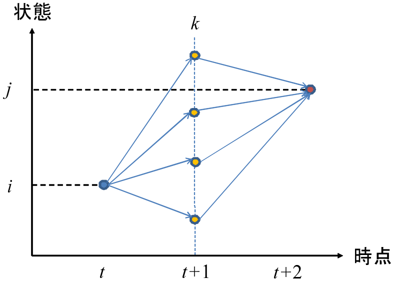


行方向の総和は $\sum_j p_{ij} = 1$

複数ステップを介した遷移確率

時点 t **で状態** i **であるとき，** m **ステップの遷移の後，時点** $t+m$ **で状態** j **にいる確率** $P(X_{t+m}=j|X_t=i)$ **について考える**

- $m=1$ **のとき，定義より** $P(X_{t+1}=j|X_t=i)=p_{ij}$
- $m=2$ **の時は** 2 **ステップで** i **から** j **へ遷移．すなわち，まず最初のステップで** i **から状態** k **に遷移し，そこから** j **へ遷移．**
- **この確率は** $P(X_{t+2}=j, X_{t+1}=k|X_t=i)=p_{ik}p_{kj}$



$m = 2$ の時の $P(X_{t+m} = j | X_t = i)$

$P(X_{t+2} = j, X_{t+1} = k | X_t = i)$ を乗法定理を使って書き直すと

$$P(X_{t+2} = j, X_{t+1} = k | X_t = i) = \\ P(X_{t+1} = k | X_t = i) P(X_{t+2} = j | X_{t+1} = k, X_t = i)$$

マルコフ性より，右辺第 2 項の $X_t = i$ は無視できるので上式は

$$P(X_{t+1} = k | X_t = i) P(X_{t+2} = j | X_{t+1} = k)$$

となる．したがって，ある状態 k を介して状態 i から j へ状態が遷移する確率は

$$P(X_{t+2} = j, X_{t+1} = k | X_t = i) = p_{ik} p_{kj}$$

となる．これを可能なすべての k について足し合わせれば

$$P(X_{t+2} = j | X_t = i) = \sum_k p_{ik} p_{kj} = p_{ij}^{(2)}$$

m ステップを介した遷移確率の一般化

- $m=3$ のとき，2 ステップで i から k に遷移し，そこから 1 ステップで j に行くと考えれば，同様にして $P(X_{t+3}=j|X_t=i)=\sum_k p_{ik}^{(2)} p_{kj}$
- 一般に m のとき，$P(X_{t+m}=j|X_t=i)=\sum_k p_{ik}^{(m-1)} p_{kj}$
- 遷移確率行列 P を用いると $\boldsymbol{P}^{(m)}=\boldsymbol{P}^{(m-1)}\boldsymbol{P}$．ここで $\boldsymbol{P}^{(m)}$ の第 ij 要素は $p_{ij}^{(m)}$

【例】今日が晴のとき，2 日後に晴である確率を求めてみよう．

- 晴 $\rightarrow$ 晴 $\rightarrow$ 晴 (0.6×0.6)
- 晴 $\rightarrow$ 曇 $\rightarrow$ 晴 (0.3×0.3)
- 晴 $\rightarrow$ 雨 $\rightarrow$ 晴 (0.1×0.2)

したがって，$0.36+0.09+0.02=0.47$

時点 t における状態確率分布 (state probability distribution)

時刻 t で状態 i $(i = 1, ..., N)$ にいる確率を $q_i(t) = P(X_t = i)$ とする ($\sum_{i=1}^{N} q_i(t) = 1$)

時点 t における状態確率分布

$$\boldsymbol{\pi}(t) = (q_1(t), q_2(t), ..., q_N(t))$$

とくに $t = 0$ のときの分布 $\pi(0)$ を初期分布と呼ぶ．これを用いれば，確率 $P(X_0 = i, X_t = j)$ は初期確率 $q_i(0)$ と t 時の遷移確率 $p_{ij}^{(t)}$ との積で表される．これらを全ての i について加えると

$$q_j(t) = \sum_{i=1}^{N} q_i(0) p_{ij}^{(t)}$$

時点 t における状態確率分布の続き

時点 t における分布は

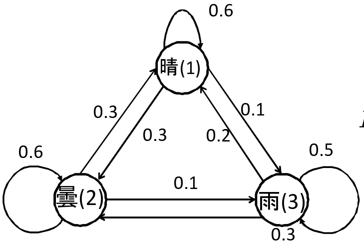
$$\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0)\boldsymbol{P}^{(t)} = \boldsymbol{\pi}(t-1)\boldsymbol{P}$$

ここで

$$\boldsymbol{\pi}(0)\boldsymbol{P}^{(t)} = \boldsymbol{\pi}(0)\boldsymbol{P}^{(t-1)}\boldsymbol{P} = \boldsymbol{\pi}(t-1)\boldsymbol{P}$$

すなわち，$\pi(0)$ に P を繰り返し掛けることにより $\pi(t)$ を順次求める．
これにより t 日後の天気が晴，曇，雨，となる確率を簡単に求めることができる．

例題

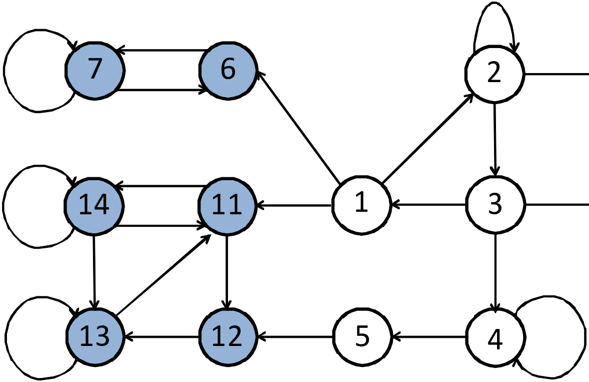


$$P = \begin{pmatrix} 0.6 & 0.3 & 0.1 \\ 0.3 & 0.6 & 0.1 \\ 0.2 & 0.3 & 0.5 \end{pmatrix}$$

- **今日の天気が晴だとする**: $\pi(0) = (1, 0, 0)$
- **明日の天気の分布**: $\pi(1) = \pi(0)P = (0.6, 0.3, 0.1)$
- **明後日の天気の分布**: $\pi(2) = \pi(1)P = (0.47, 0.39, 0.14)$

マルコフ連鎖の様々な型

矢印はすべて正の遷移確率があるとする



既約 (吸収的) な集合

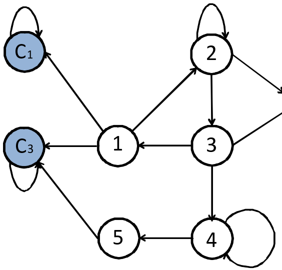
状態 1，2，3 の間を回っているうち $C_1 = \{6, 7\}$，$C_2 = \{8, 9, 10\}$，$C_3 = \{11, 12, 13, 14\}$ の三つの集合のうちの一つに入ると，その中だけで遷移するようになる．このような集合を既約な (吸収的な) 集合と呼ぶ．既約な集合 C は以下の条件で特徴づけられる

- C の中のどの二つの状態をとっても一方から他方へ何ステップかで遷移できる
- C の中のどの状態からも C の外への状態へ遷移できない

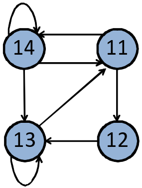
一般のマルコフ連鎖の状態空間 (取り得る状態全体からなる集合) を S，R 個の既約な集合を C_r $(r = 1, ..., R)$，それ以外の状態の集合を T とすれば $S = T \cap C_1 \cap C_2 \cdots \cap C_R$ と書ける

吸収的マルコフ連鎖，既約なマルコフ連鎖

- C_r に入るまでの遷移の様子を調べる場合は各 C_r をまとめたマルコフ連鎖を調べる
- 各 C_r の遷移の様子を調べるには，C_r を状態空間とするマルコフ連鎖を考える



既約な集合をまとめて作った
吸収的マルコフ連鎖



C3 を状態空間とする
既約なマルコフ連鎖

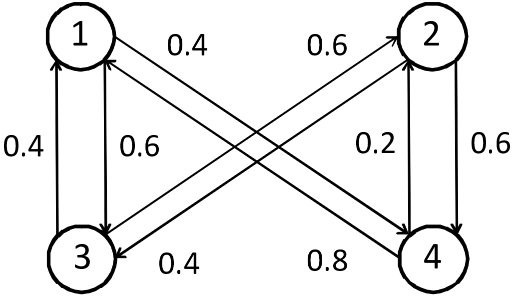
既約で非周期的なマルコフ連鎖の例

天気の移り変わりの例は非周期的で既約なマルコフ連鎖

- $\pi(0) = (1, 0, 0)$
- $\pi(1) = (0.6, 0.3, 0.1)$
- $\pi(2) = (0.47, 0.39, 0.14)$
- $\pi(3) \sim (0.43, 0.42, 0.15)$
- $\pi(4) \sim (0.41, 0.43, 0.16)$
- $\pi(5) \sim (0.41, 0.43, 0.16)$
- $\pi(6) \sim (0.41, 0.43, 0.16)$
- $\pi(7) \sim (0.41, 0.43, 0.16)$

初期分布の影響が毎日の天気の移り変わりの偶然性によって減少し，4 日目よりその影響はほとんど無視される

既約で周期的なマルコフ連鎖の例



- 状態 1 から出発すると，偶数時点では状態 1 または 2 にいる．一方，奇数時点では状態 3 または 4 にいる
- 各 C_r の遷移の様子を調べるには，C_r を状態空間とするマルコフ連鎖を考える

各時点の分布

$$
\boldsymbol{P} = \begin{pmatrix} 0 & 0 & 0.6 & 0.4 \\ 0 & 0 & 0.4 & 0.6 \\ 0.4 & 0.6 & 0 & 0 \\ 0.6 & 0.4 & 0 & 0 \end{pmatrix}
$$

- $\boldsymbol{\pi}(0) = (1, 0, 0, 0)$
- $\boldsymbol{\pi}(1) = (0, 0, 0.6, 0.4)$
- $\boldsymbol{\pi}(2) = (0.56, 0.44, 0, 0)$
- $\boldsymbol{\pi}(3) = (0, 0, 0.51, 0.49)$
- $\boldsymbol{\pi}(4) = (0.60, 0.40, 0, 0)$
- $\boldsymbol{\pi}(5) = (0, 0, 0.52, 0.48)$
- $\boldsymbol{\pi}(6) = (0.59, 0.41, 0, 0)$
- $\boldsymbol{\pi}(7) = (0, 0, 0.52, 0.48)$

$\boldsymbol{\pi}(t) \neq \boldsymbol{\pi}(0)\boldsymbol{P}^{(t)}$ **に注意！**

エルゴード的マルコフ連鎖

エルゴード性 (ergodic property)

- **どの状態から出発しても，どの状態にも遷移する可能性がある** (**既約**)
- **周期性を持たない**

エルゴード性を持ったマルコフ連鎖をエルゴード的マルコフ連鎖 (ergodic Markov chain) **とよ呼ぶ．情報源と通信路で出てくるので覚えておくこと．**

定常分布 (stationary distribution)

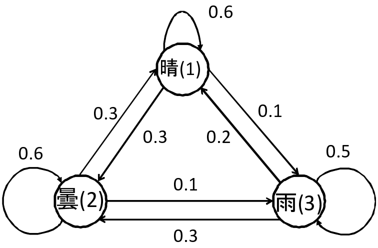
時間が経過しても確率過程の分布が変化しない場合，その分布を定常分布と呼ぶ

【天気の移り変わりの例】

- **逐次計算では $\pi(t)$ を順次計算して定常分布を求める**:
  $\boldsymbol{\alpha} \sim \boldsymbol{\pi}(10) = (0.4048, 0.4286, 0.1666)$
- **遷移方程式では α が唯一の確率ベクトルである性質を用いる．すなわち，求めたい定常分布を $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \alpha_3)$ として $\boldsymbol{\alpha} = \boldsymbol{\alpha} P$ を解けば良い**:
- $\alpha_1 = 0.6\alpha_1 + 0.3\alpha_2 + 0.2\alpha_3$
- $\alpha_2 = 0.3\alpha_1 + 0.6\alpha_2 + 0.3\alpha_3$
- $\alpha_3 = 0.1\alpha_1 + 0.1\alpha_2 + 0.5\alpha_3$
  これらと $\alpha_1 + \alpha_2 + \alpha_3 = 1$ を用いて解く

平衡方程式による定常分布の計算

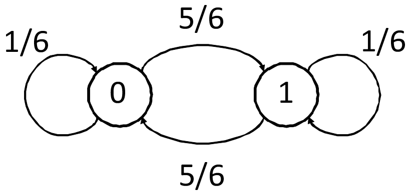
状態遷移確率を荷物に例えて，ある状態から他の状態へ移る量と，入ってくる量が等しいとして解く



- **晴の状態から出て行く量**: $(0.3 + 0.1)\alpha_1$
- **晴の状態に入ってくる量**: $0.3\alpha_2 + 0.2\alpha_3$
- **平衡状態ではこれらが等しいので** $0.4\alpha_1 = 0.3\alpha_2 + 0.2\alpha_3$
- **他の状態でも同様に方程式を作成して解く**

問題

以下のエルゴード的マルコフ連鎖について答えよ．



- 【4.1】0 を状態 1，1 を状態 2 として遷移確率行列 P を書け
- 【4.2】$\pi(0) = (1, 0)$ のとき $\pi(1)$ を求めよ
- 【4.3】定常分布を求めよ