マルコフ連鎖

樋口さぶろお https://hig3.net

龍谷大学大学院 理工学研究科 数理情報学専攻

理論物理学特論　L06 (2022-10-26 Wed)

最終更新: Time-stamp: "2022-10-26 Wed 08:31 JST hig"

今日の目標

- 分布のパラメタを確率変数とみなして事後分布を求められる
- マルコフ連鎖の定義と例を説明できる

略解 I

L05-Q1

$$\Theta \sim \mathrm{N}(-2, 3^2) \rightsquigarrow \Theta | Y = -3 \sim \mathrm{N}(-\tfrac{35}{13}, \tfrac{36}{13})$$

L05-Q2
略

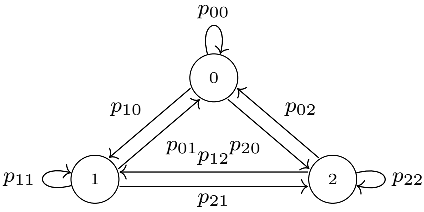
ここまで来たよ

5 ベイズ推定

6 マルコフ連鎖
- マルコフ連鎖
- 確率ベクトル, 転置確率行列
- 定常分布, 分布の時間発展

マルコフ連鎖の例 I

時間 t と共に変化する猫の状態
0: 食べる 1: 寝る 2: 遊ぶ を考える.
状態空間 $S = \{0, 1, 2\}$
状態 $x = 0$: 食べる, ...
時刻 t の猫の状態 $X(t) = 0, 1, 2$.



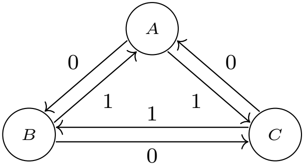
推移確率

条件付き確率

$$p_{先\ 元} = p_{xy} = P(X(t) = x | X(t-1) = y)$$

上の量を, 世の中では p_{xy} でなく p_{yx} と書くこともある.
マルコフ連鎖「このような」状態空間と, 推移確率の組.

比較: オートマトンの状態遷移図では, 矢印の数値は確率でなく入力



$p(x,t)$ の漸化式

略記 $p(x,t) = P(X(t) = x)$.
条件付き確率の全確率の法則から,

$$
\begin{aligned}
p(0,t) =& p_{00} \cdot p(0,t-1) + p_{01} \cdot p(1,t-1) + p_{02} \cdot p(2,t-1) \\
p(1,t) =& p_{10} \cdot p(0,t-1) + p_{11} \cdot p(1,t-1) + p_{12} \cdot p(2,t-1) \\
p(2,t) =& p_{20} \cdot p(0,t-1) + p_{21} \cdot p(1,t-1) + p_{22} \cdot p(2,t-1)
\end{aligned}
$$

$$
\begin{pmatrix} p(0,t) \\ p(1,t) \\ p(2,t) \end{pmatrix} = \begin{pmatrix} p_{00} & p_{01} & p_{02} \\ p_{10} & p_{11} & p_{12} \\ p_{20} & p_{21} & p_{22} \end{pmatrix} \begin{pmatrix} p(0,t-1) \\ p(1,t-1) \\ p(2,t-1) \end{pmatrix}
$$

$$
p(x,t) = \sum_{y=0,1,2} p_{xy} \cdot p(y,t-1). \qquad (x = 0,1,2)
$$

転置推移確率行列 M, 分布=ベクトル $\vec{p}(t)$ で書くと (x, y が成分番号)

$$
\vec{p}(t) = M\vec{p}(t-1).
$$

この漸化式を解くと, $\vec{p}(t) = M\vec{p}(t-1) = M^2\vec{p}(t-2) = \cdots = M^t\vec{p}(0)$.

L06-Q1

Quiz(マルコフ連鎖の転置推移確率行列)

状態 $x = 0, 1, 2, 3$ からなる状態空間を持つマルコフ連鎖を考える．時刻 $t = 0$ に状態は $x = 1$ である．時刻 t に状態が x であるという条件のもとで，時刻 $t + 1$ での状態は，

- 確率 $\frac{1}{7}$ で状態 $x - 1$,
- 確率 $\frac{2}{7}$ で状態 $x + 1$,
- 確率 $\frac{4}{7}$ で状態 x.

ただし，上の条件付き確率で，状態 $x = -1$ とは状態 $x = 3$ のこと，状態 $x = 4$ とは状態 $x = 0$ のことをいう．

1. 推移図を書こう．
2. 転置推移確率行列を書こう．
3. 時刻 $t = 0$ における分布 $\vec{p}(0) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ であるとき，$\vec{p}(1)$ を求めよう．

L06-Q2

Quiz(マルコフ連鎖の転置推移確率行列)

状態 $x = 0, 1, 2, 3$ からなる状態空間を持つマルコフ連鎖を考える．時刻 $t = 0$ に状態は $x = 1$ である．時刻 t に状態が x であるという条件のもとで，時刻 $t + 1$ での状態は，

- 確率 $\frac{1}{7}$ で状態 $x - 1$,
- 確率 $\frac{2}{7}$ で状態 $x + 1$,
- 確率 $\frac{4}{7}$ で状態 x.

ただし，上の条件付き確率で，状態 $x = -1$ とは状態 $x = 0$ のこと，状態 $x = 4$ とは状態 $x = 3$ のことをいう．

1. 推移図を書こう．
2. 転置推移確率行列を書こう．
3. 時刻 $t = 0$ における分布 $\vec{p}(0) = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ であるとき,$\vec{p}(1)$ を求めよう．

ここまで来たよ

5 ベイズ推定

6 マルコフ連鎖
- マルコフ連鎖
- 確率ベクトル, 転置確率行列
- 定常分布, 分布の時間発展

線形代数のりで確率ベクトル, 転置確率行列の言葉を準備 I

定義 (非負ベクトル)

m 次元ベクトル $\begin{pmatrix} p_0 \\ p_1 \\ \vdots \\ p_{m-1} \end{pmatrix}$ が $p_i \geq 0$ を満たすとき, 非負ベクトルという.

定義 (確率ベクトル)

非負ベクトル $\begin{pmatrix} p_0 \\ p_1 \\ \vdots \\ p_{m-1} \end{pmatrix}$ が, 規格化 $\sum_{i=0}^{m-1} p_i = 1$ を満たすとき, 確率ベクトルという.

線形代数のりで確率ベクトル, 転置確率行列の言葉を準備 II

離散型確率分布 $f(x)$ は, 確率ベクトルと 1 対 1 に対応.

$$\vec{p} = \begin{pmatrix} p_0 \\ p_1 \\ p_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} f(0) \\ f(1) \\ f(2) \\ \vdots \end{pmatrix}$$

定義 (非負行列 ($n \times m = 3 \times 2$ の例で))

実行列 $\begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \\ p_{20} & p_{21} \end{pmatrix}$ が $p_{ij} \geq 0$ を満たすとき**非負行列**という.

定義 (転置確率行列)

行列 $M = \begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \\ p_{20} & p_{21} \end{pmatrix}$ の各列ベクトルが確率ベクトルであるとき, つまり $\forall j \sum_{i=0}^{n-1} p_{ij} = 1$ であるとき, M を**転置確率行列**という.

転置推移確率行列

定義 (転置推移確率行列)

マルコフ連鎖に現れる，転置確率行列

$$M = \begin{pmatrix} p_{00} & p_{01} & p_{02} \\ p_{10} & p_{11} & p_{12} \\ p_{20} & p_{21} & p_{22} \end{pmatrix}$$

を，マルコフ連鎖の**転置推移確率行列**という．$m \times m$ 正方行列．

世の中では，それぞれの転置で (${}^{\mathrm{t}}M$: **推移確率行列**) 次のようにに書くのがふつう．

$${}^{\mathrm{t}}\vec{p}(t) = {}^{\mathrm{t}}\vec{p}(t-1)\ {}^{\mathrm{t}}M$$

漸化式 $\vec{p}(t) = M\vec{p}(t-1)$ の解

$$\vec{p}(t) = M\vec{p}(t-1) = M^2\vec{p}(t-2) = \cdots = M^t\vec{p}(0).$$

命題 (転置確率行列と確率ベクトルの積)

M が転置確率行列, $\vec{p}$ が確率ベクトルのとき, $M\vec{p}$ も確率ベクトル

証明:

確率過程とマルコフ連鎖

確率過程　時間 t に依存する確率変数

もちろん反復試行から来る独立な確率変数群 $R(t)$ も時間 t に依存するけど, 独立じゃない場合がおもしろい

離散時間マルコフ連鎖は, 次の性質を満たす確率過程.

- 離散時間　t が離散的. $t = 0, 1, 2, \ldots$
- マルコフ Markov　$\vec{p}(t)$ が直前の時刻の分布 $\vec{p}(t-1)$ だけから決まる. 転置推移確率行列 p_{xy} で表現できる.
- 連鎖 chain　状態空間 $S \ni x$ が離散的.

いま考えてる, 時間空間離散のランダムウォークや猫は離散時間マルコフ連鎖の典型例.

離散時間マルコフ連鎖の (確率) 分布は, 確率ベクトルで表現できる.

離散時間マルコフ連鎖の遷移確率は, 転置確率行列で表現できる.

ここまで来たよ

5 ベイズ推定

6 マルコフ連鎖
- マルコフ連鎖
- 確率ベクトル, 転置確率行列
- 定常分布, 分布の時間発展

定義 (定常分布)

$\vec{p} = M\vec{p}$ となる分布 $\vec{p}$ を定常分布という.

意味 自分の言葉でどうぞ

線形代数の言葉で言うと, 定常分布は, 転置推移確率行列 M の

自分の言葉でどうぞ

建設的心配性大爆発

- 定常分布っていつでもある? $\rightsquigarrow$ Yes
- 固有値 (の絶対値) が 1 より大きな固有ベクトルはあるの? $\rightsquigarrow$ No

状態数 m が有限のとき, ペロン・フロベニウスの定理から言える.

命題 (固有値 1 の存在)

転置確率行列は 固有値 1 を持つ.

証明

分布の時間発展 I

L06-Q3

Quiz(マルコフ連鎖の時間発展)

状態空間 $\{0,1\}$ 上のマルコフ連鎖を考える. 転置推移確率行列は

$$M = \begin{pmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{2} \end{pmatrix}.$$

1. 初期分布 $\vec{p}(0) = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ のとき $\vec{p}(1)$ を求めよう.
2. 定常分布をすべて求めよう.
3. 上の初期分布のとき $\vec{p}(t)$ を求めよう.
4. 上の初期分布のとき極限分布 $\vec{p}(\infty) = \lim_{t\to+\infty} \vec{p}(t)$ と収束の様子 $\log|\vec{p}(t) - \vec{p}(\infty)|$ を調べよう.
5. 初期分布 $\vec{p}(0) = \frac{1}{2}\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ のとき $\vec{p}(t)$ を求めよう.

Quiz(マルコフ連鎖の定常分布)

次の転置推移確率行列を持つ 状態空間 $\{0,1,2\}$ 上のマルコフ連鎖を考えよう.

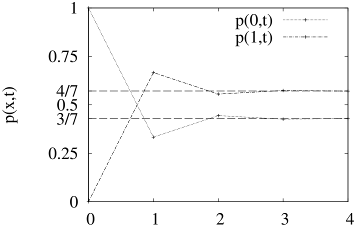
$$M = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{2}{4} & \frac{1}{4} \\ 0 & \frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

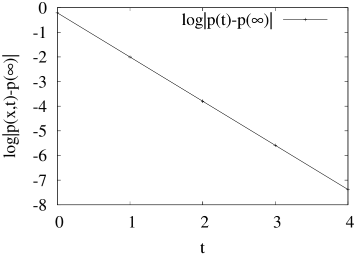
なお, M の固有値固有ベクトルは $\lambda = 1, \frac{3}{4}, \frac{1}{4}, \quad \vec{u} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ であることを使ってよい.

1. 定常分布をすべて求めよう.
2. 初期分布 $\vec{p}(0) = \frac{1}{3}\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$ のとき $\vec{p}(t)$ を求めよう.
3. 上のとき, $\log|\vec{p}(t) - \vec{p}(\infty)|$ の $t \to +\infty$ の振る舞いを求めよう.

極限分布

(存在するなら) $\vec{p}(+\infty) = \lim_{t\to+\infty} \vec{p}(t)$ を極限分布という.




log|p(t)-p(∞)|
log|p(x,t)-p(∞)|
0
-1
-2
-3
-4
-5
-6
-7
-8
0
1
2
3
4
t

単純な場合でのマルコフ連鎖の時間発展 I

状態空間 $S=\{0,1,\ldots,m-1\}$ 上のマルコフ連鎖.
転置推移確率行列 M の固有値固有ベクトル $\lambda_i, \vec{u}_i$ $(i=1,\ldots,m)$.

仮定

絶対値の大きさの順に並べると (第 1, 第 2,. . ., 第 m 固有値とよぶ) 次のようになっているとする.

$$1=\lambda_1>|\lambda_2|\geq\cdots\geq|\lambda_m|\geq 0.$$

このとき, $\vec{u}_1$ は確率ベクトルにとれて, 解は,

$$\vec{p}(t)=\vec{u}_1+a_2\vec{u}_2\lambda_2^t+a_3\vec{u}_3\lambda_3^t+\cdots+a_m\vec{u}_m\lambda_m^t.$$

$$\vec{p}(t) = \vec{u}_1 + a_2 \vec{u}_2 \lambda_2^t + a_3 \vec{u}_3 \lambda_3^t + \cdots + a_m \vec{u}_m \lambda_m^t.$$

この単純な場合の性質

- 第 1 固有値は 1. 転置推移確率行列の固有値には, いつでも 1 が含まれることを示したのだった. 先週の証明
- 第 1 固有ベクトル $\vec{u}_1$ は 定常分布. ペロン-フロベニウスの定理
- 極限分布 $\vec{p}(+\infty) = \lim_{t \to +\infty} \vec{p}(t) = \vec{u}_1$. 自分の言葉でどうぞ

ちなみに, 極限分布が存在するなら, それは必ず定常分布. なぜなら 自分の言葉でどうぞ

この単純な場合の性質の続き

- 第 2 以降の固有値の絶対値が 1 より小なので 自分の言葉でどうぞ
  第 2 固有値の絶対値が小ささが, 極限分布への収束の速さを決める
- 第 2 以降の固有ベクトルは 自分の言葉でどうぞ

これは単純な場合で, 一般にはそうでない場合もある

- 単純である＝既約 (可約でない) かつ非周期的
- 単純でない一般的な場合＝可約または周期的