マルコフ連鎖と定常分布

Cubic S

1. マルコフ連鎖

1.1. マルコフ連鎖とは

マルコフ連鎖とは、確率変数列$\{\mathrm{X_t}\}$で$\mathrm{X_{n+1}}$の条件付き確率分布が$\mathrm{X_n}$にのみ依存するもののことをいう。この条件を数式で表すと以下のようになる。

$$\Pr(\mathrm{X_{n+1}} = x \mid \mathrm{X_0} = x_0, \mathrm{X_1} = x_1, \cdots, \mathrm{X_n} = x_n) = \Pr(\mathrm{X_{n+1}} = x \mid \mathrm{X_n} = x_n)$$

「1 秒後の状態遷移は現在の状態にのみ依存する」、というのを数学的に定式化したものである。

特に、マルコフ連鎖$\{\mathrm{X_t}\}$のとりうる値が有限個の場合、**有限状態マルコフ連鎖**といい、その値のことを**状態**という。また、全ての時間 t に対して

$$\Pr(\mathrm{X_{t+1}} = x \mid \mathrm{X_t} = y) = \Pr(\mathrm{X_1} = x \mid \mathrm{X_0} = y)$$

が成り立つとき、**時間的に均一なマルコフ連鎖**という。

時間的に均一な有限状態マルコフ連鎖$\{\mathrm{X_t}\}$において、状態$x_0, x_1, \cdots x_n$として、第(i, j)成分$p_{i,j}$を

$$p_{i,j} = \Pr(\mathrm{X_{n+1}} = x_j \mid \mathrm{X_n} = x_i)$$

とした行列 P を$\{\mathrm{X_t}\}$における**状態遷移行列**という。

1.2. 基本性質

マルコフ連鎖のいくつかの基本性質について説明する。以下$\{\mathrm{X}_{\mathrm{t}}\}$を時間的に均一な有限状態マルコフ連鎖とする

1.2.1. n個先の確率

$\{\mathrm{X}_{\mathrm{t}}\}$に対する状態遷移行列をPとすると、現時点の状態が$x_{\mathrm{i}}$のときの$n$個先の未来の状態が$x_{\mathrm{j}}$である確率は、行列の積を使って求めることが出来る。

$$
\begin{aligned}
&\Pr(\mathrm{X}_{\mathrm{t}+n} = x_j \mid \mathrm{X}_{\mathrm{t}} = x_i) \\
&= \sum\nolimits_k \Pr(\mathrm{X}_{\mathrm{t}+n} = x_j \mid \mathrm{X}_{\mathrm{t}+n-1} = x_k)\Pr(\mathrm{X}_{\mathrm{t}+n-1} = x_k \mid \mathrm{X}_{\mathrm{t}} = x_i) \\
&= \sum\nolimits_k p_{i,j}\Pr(\mathrm{X}_{\mathrm{t}+n-1} = x_k \mid \mathrm{X}_{\mathrm{t}} = x_i)
\end{aligned}
$$

ここで、行ベクトルV_{i} を$\left(\Pr(\mathrm{X}_{\mathrm{i}} = x_1), \Pr(\mathrm{X}_{\mathrm{i}} = x_2), \cdots, \Pr(\mathrm{X}_{\mathrm{i}} = x_n)\right)$とすると、

$$\Pr(\mathrm{X}_{\mathrm{t}+n} = x_j \mid \mathrm{X}_{\mathrm{t}} = x_i) = \mathrm{V}_{\mathrm{t}+n\text{-}1} * \mathrm{P}$$

となる。この操作を$\mathrm{V}_{\mathrm{t}+n\text{-}1}$について$n-1$回繰り返すと、

$$\Pr(\mathrm{X}_{\mathrm{t}+n} = x_j \mid \mathrm{X}_{\mathrm{t}} = x_i) = \mathrm{V}_{\mathrm{t}} * \mathrm{P}^n$$

となる。

1.2.2. 可約性

任意の状態x_i, x_jに対して、

$$\Pr(\mathrm{X}_n = x_j \mid \mathrm{X}_0 = x_i) > 0$$

を満たす $n > 0$ が存在するとき、**既約**なマルコフ連鎖という。

1.2.3. 周期性

任意の状態iに対して、

$$k = \gcd\{n \mid \Pr(\mathrm{X}_n = i \mid \mathrm{X}_0 = i) > 0\}$$

を状態iの**周期**といい、特に$k = 1$のとき状態iは**非周期的**であるという。

既約かつすべての状態が非周期的であるときマルコフ連鎖は**エルゴート的**であるという

2. 定常分布

2.1. 定常分布の定義

確率変数Yが、時間的に均一な有限状態マルコフ連鎖$\{\mathrm{X_t}\}$と全ての状態x_iに対し、

$$\Pr(\mathrm{Y}=x_i)=\sum_j \Pr(\mathrm{Y}=x_j)p_{j,i}$$

が成り立つとき、Yを$\{\mathrm{X_t}\}$における**定常分布**という。

2.2. 定常分布の存在

時間的に均一な有限状態マルコフ連鎖で、全ての$p_{i,j}$で正ならば定常分布

は常に存在する。
全てのjに対して

$$\sum_j p_{i,j}=1$$

が成り立つので、
P‐E　　(E は単位行列)
の行列式は 0 となる。よって、
V*(P‐E) = 0
となる行ベクトル V が存在する。
この V の成分はすべて正かすべて負かのいずれかであることを示す。もし、V の成分で、$v_i>0, v_j<0$となるものが存在したとすると、

$$\sum_{j,v_j<0} v_j=\sum_{l,j,v_j<0} v_l p_{l,j} > \sum_{l,j,v_j<0,v_l<0} v_l p_{l,j} > \left(\max_{i,j} p_{i,j}\right)\sum_{l,v_l<0} v_l$$

が成り立つ。$\left(\max_{i,j} p_{i,j}\right)<1$かつ$\sum_{l,v_l<0} v_l<0$となり矛盾が生じる。したがって V の成分はすべて正かすべて負かのいずれかである。この V の成分が非負、かつ和が 1 になるように定数倍したものが定常分布になる。

2.3. 極限分布と定常状態

時間的に均一な有限状態マルコフ連鎖で、既約かつ非周期的でならば、任意の行ベクトル V(全ての成分が非負で和が 1 になる)に対して、

$\mathrm{V*P}^k \;\to\; \pi \quad (k \to \infty)$

が成り立つ。(π は定常分布)

既約性と非周期性から $p_{i,j} > 0$ としてよい(証明は Appendix に掲載)

$\mathrm{V}_1 = \mathrm{V*P}$ とすると、V に依存しないある定数 $0 < c < 1$、全ての成分が非負で和が 1 になる行ベクトル R_1 を用いて

$\mathrm{V}_1 = c\pi + (1-c)\mathrm{R}_1$

と書き表せる。$\mathrm{V}_2 = \mathrm{V}_1\mathrm{*P}, \mathrm{V}_3 = \mathrm{V}_2\mathrm{*P}, \cdots, \mathrm{V}_k = \mathrm{V}_{k-1}\mathrm{*P}$ とすると、上記と同じく、全ての成分が非負で和が 1 になる行ベクトル R_k に対して

$$\mathrm{V}_k = \{1-(1-c)^k\}\pi + (1-c)^k \mathrm{R}_k$$

とできる。$0 < 1-c < 1, \;\; |\mathrm{R}_k| \leq 1$ より $k \to \infty$ とすると、$\mathrm{V*P}^k \to \pi \quad (k \to \infty)$

が成り立つ。

3. マルコフ連鎖と定常分布の性質を用いた例

この章では、マルコフ連鎖の考えを用いた乱数生成のアルゴリズム「メトロポリス法」を紹介する。

3.1. マルコフ連鎖の詳細つりあい条件

時間的に均一な有限状態マルコフ連鎖$\{\mathrm{X}_t\}$と、確率分布$\pi(x)$に対して、以下の条件が成り立つとき**詳細つりあい条件を満たす**という。

『任意の状態 x_i, x_j に対して

$$\pi(x_i)p_{i,j} = \pi(x_j)p_{j,i}$$

が成り立つ』

ただし、$p_{i,j}$ は$\{\mathrm{X}_t\}$に対する状態遷移行列 P の第(i, j)成分である。

これは、『π が$\{\mathrm{X}_t\}$に対する定常分布になる』という命題の十分条件になる。この条件を満たすものに限定すれば、与えられた確率分布 π に対するマルコフ連鎖が設計できる。

3.2. メトロポリス法

メトロポロス法は、『生成する状態の「候補」を作り、採択するかどうかを確率的に決める』というプロセスにより、望みの分布π（に限りなく近い分布）に従う乱数を生成方法である。メトロポリス法の手順は以下の通りである。

1. 現在の状態を x_t としたとき、確率分布 $f(y)=Q(x_t,y)$ に従う候補 y_t 生成する。
2. $r = \pi(y_t) \big/ \pi(x_t)$ を計算する。
3. $0 \le R < 1$ の一様乱数 R を生成する。
4. $R < r$ なら「候補」を次の状態 x_{t+1} として採用する。そうでなければ x_{t+1} を x_t とする。

ただし、$Q(x_i,x_j)$ は状態 x_i であるときに x_j が候補として選ばれる確率で、エルゴート的であり、かつ $Q(x_i,x_j)=Q(x_j,x_i)$ が成り立つものとする。また、全ての x_i について $\pi(x_i)>0$ が成り立つものとする。

この方法でπに限りなく近い分布に従う状態を生成できることを説明する。

状態遷移行列 P の第 (i, j) 成分が、

$$p_{i,j} = \begin{cases} Q(x_i,x_j)+\alpha_i & if \quad i=j \\ Q(x_i,x_j) & if \quad i \ne j \quad and \quad \pi(x_i) \le \pi(x_j) \\ Q(x_i,x_j)\pi(x_j) \big/ \pi(x_i) & otherwise \end{cases}$$

となるようなマルコフ連鎖 $\{X_t^{'}\}$ を考えると、$\{X_t^{'}\}$ とπについて詳細つりあい条件が成り立つ。ただし、α_i は、$\sum_j p_{i,j} = 1$ となるように定められた正の定数である。

$Q(x_i,x_j)$ の既約性から、全ての x_i, x_j に対して

$$\overbrace{\sum\sum\cdots\sum}^{m} Q(x_i,x_l)Q(x_l,x_k)\cdots Q(x_w,x_j) > 0$$

となるmが存在する。これより$Q(x_i, x_l)Q(x_l, x_k)\cdots Q(x_w, x_j) > 0$となるものが存在

する。P^mの第(i, j)成分は、1 以下の正の定数$m_{l,k,\cdots w}$と非負の定数βを用いて

$$\beta + \overbrace{\sum\sum\cdots\sum}^{m} m_{l,k,\cdots w} Q(x_i, x_l)Q(x_l, x_k)\cdots Q(x_w, x_j)$$

となるので、$\{\mathrm{X}_\mathrm{t}^{'}\}$は既約である。

同じ方法で、$Q(x_i, x_j)$の非周期性から$\{\mathrm{X}_\mathrm{t}^{'}\}$の非周期性も証明できる。

これより、2.3.の事実から上記の手順で生成された状態x_tは、tが十分大ならばπに限りなく近い分布に従うことが分かる。

Appendix

ここでは、前回証明を飛ばした「時間的に均一な有限状態マルコフ連鎖$\{\mathrm{X_t}\}$が既約かつ全ての状態について非周期的ならば、ある$M>0$で、『すべての$m \geq M, i, j$に対してP^mの第(i,j)成分が正である』が成り立つものが存在する」という命題を証明する。

まず、自然数の部分集合Aで

(1) gcd A=1

(2) Aに属する自然数a, a'に対して、$a+a'$はAに属する

という条件を満たすならば、ある$M>0$で、『すべての$m \geq M$はAに属する』が成り立つものが存在することを示す。

一番目の条件よりAに属するある自然数$a_1, a_2, \cdots, a_n$と、ある整数$m_1, m_2, \cdots, m_n$に対し、

$$\sum a_i m_i = 1$$

が成り立つ（ユークリッドの互除法により証明できる）m_iで負となるものを$m^{neg}{}_i$，そうでないものを$m^{pos}{}_i$と置くと、二番目の条件より

$$\sum a_i(-m^{neg}{}_i)+1=\sum a_i m^{pos}{}_i$$

から$\sum a_i(-m^{neg}{}_i)+1$は、Aに属する。これより、$l, l+1$がAに属するような自然数が存在することが分かる。ここで、$M \equiv l^2$とおく。このとき、$m \geq M$となる自然数m に対して、

$$L\cdot l \leq m < L\cdot(l+1)$$

となる整数$L \geq l$が存在する。ここで、$n \equiv m - L\cdot l$と

$$n(l+1)+(L-n)l = m$$

が成り立ち、これは、mがAに属する自然数であることを意味している。

次に、「時間的に均一な有限状態マルコフ連鎖$\{\mathrm{X_t}\}$が既約かつ全ての状態について非周期的ならば、ある$N>0$で、『すべての$m \geq N, i$に対してP^mの第(i,i)成分が正

である』が成り立つものが存在する」を示す。

各iに対して、$A_i \equiv \{n \mid (P^n)_{i,i} > 0\}$とおく。

このA_iについて、非周期性から前の命題の条件(1)を満たす。A_iに属する自然数a,bに対して、

$$(P^{a+b})_{i,i} \geq (P^a)_{i,i}(P^b)_{i,i} > 0$$

より、$a+b$はAに属する。これよりこの命題の証明が完了した。

最後にこの Appendix で示すべきメインの命題を示す。

各i, jに対して$(P^{m_{i,j}})_{i,i}$>0 となる$m_{i,j}$がとれる。ここで、$M \equiv m_{i,j} + N$とおくと

$m \geq N$に対して、$(P^m)_{i,j} > 0$が成り立つ。これより、$M \equiv \max_{i,j}(m_{i,j} + N)$とすれば望みの定数が得られる。

参考文献

- 伊庭幸人,種村正美,大森裕浩,和合肇,佐藤整尚,高橋明彦(2005), 統計科学のフロンティア　計算統計Ⅱ, 岩波書店
- OLLE HAGGSTROM (2002), *Finite Markov Chains and Algorithmic Applications*, Cambridge Univ. Press
- ウィキペディア　マルコフ連鎖
  http://ja.wikipedia.org/wiki/%E3%83%9E%E3%83%AB%E3%82%B3%E3%83%95%E9%80%A3%E9%8E%96
- 信州大学　確率論
  http://markun.cs.shinshu-u.ac.jp/learn/probability/i_04-00.html