マルコフ連鎖 (Markov chain)

目次

定義 1 : マルコフ連鎖 (Markov chain) ある一連の確率変数 $X_0, X_1, \ddots$ に対して

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

であるとき、マルコフ連鎖という。これは現在の時刻での状態のみによって次の状態の確率が決まることである。ここで X_i のとりうる値は、連鎖の状態空間といい、加算集合となる。

状態空間が有限であるマルコフ連鎖は有限状態マルコフ連鎖という。また、斉次的 (homogeneous) マルコフ連鎖とは任意の n に対して

$$Pr(X_{n+1} = x | X_n = y) = Pr(X_n | X_{n-1} = y)$$

が成立することである。

マルコフ連鎖はコスト付きの有向グラフで表すことができ、ノードが状態、枝のコストがある状態から別の状態へ遷移する確率となる。

定義 2 : 到達可能 (accessible) あるマルコフ連鎖に対して、現在の状態が i であり、未来のある時点で状態 j にある確率が 0 でないとき、状態 j は状態 i から到達可能であるという。つまり

$$Pr(X_n = j | X_0 = i) > 0$$

となる n が存在することである。また、これはマルコフ連鎖をグラフで表わしたとき、ノード i からノード j への路が存在することと同値である。

定義 3 : 連結 (communicate)・連結類 (communicate class) あるマルコフ連鎖に対し、状態 i と状態 j が互いに到達可能ならば、状態 i と状態 j は連結しているという。そして、状態の集合 C の任意の i, j が互いに連結ならば、C は連結類といい、連結類の状態からその連結類に含まれない状態へと遷移する確率が 0 ならば、この連結類は閉じている (closed) という。

定義 4 : 既約 (または可約でない : irreducible) あるマルコフ連鎖の状態空間が連結類ならば、マルコフ連鎖は既約であるという。つまり、既約なマルコフ連鎖では任意の状態から別の任意の状態へと遷移することができる。

定義 5 : 周期

$$k = gcd\{n : Pr(X_n = i|X_0 = i) > 0\}$$

として書かれる k を状態 i の周期という。$k = 1$ のとき状態は非周期的であるという。連結類の各状態の周期は等しく、既約なマルコフ連鎖の状態が非周期的なとき、エルゴード的 (ergodic) という。

定義 6 : 遷移行列 (transition matrix) 有限状態マルコフ連鎖であれば、遷移確率分布は行列で表すことができる。このとき、遷移確率分布を表す行列 P を遷移行列 (確率行列) といい、第 (i, j) 要素 p_{ij} は

$$p_{ij} = PR(X_{n+1} = j|X_n = i)$$

と書かれる。また、斉次的なマルコフ連鎖であれば k-段階遷移確率は遷移行列の k 乗で表すことができる。

定義 7 : 定常状態 現在ある状態 i にある確率を π_i として表されるベクトル π を確率行列といい、斉次的な有限状態マルコフ連鎖の遷移行列 P に対して

$$\pi = \pi P$$

であるとき π を定常分布という。定常分布は遷移行列の左側固有ベクトルであり、固有値は 1 であるといえる。

定義 8 : ゲルシュゴリン円盤 (Gershgorin disc) $n \times n$ の複素行列 A の各成分を a_{ij} とする。この行列に対して i 行の非対角成分の絶対和 R_i を

$$R_i = \sum_{j \neq i} |a_{ij}|$$

とする。このとき、a_{ii} を中心とする半径 R_i の閉円盤 $D(a_{ii}, R_i)$ をゲルシュゴリン円盤 (Gershgorin disc) という。

定理 1 : ゲルシュゴリンの定理 (Gershgorin circle theorem) $n \times n$ の複素行列 A の任意の固有値は、少なくともひとつのゲルシュゴリン円盤 $D(a_{ii}, R_i)$ の上に乗る。

Proof. A のある固有値 λ とそれに対応する固有ベクトル $\mathbf{x}$ を一つ選び、固有ベクトル $\mathbf{x}$ の各成分のうち絶対値が最大となる成分の番号を i とする。つまり、任意の j に対して $|x_i| \geq |x_j|$

であり、$\mathbf{x}$ が固有ベクトルであることから $|x_i| > 0$ である。$\mathbf{x}$ は λ に対応する固有ベクトルとしたため $A\mathbf{x} = \lambda\mathbf{x}$ が成立し、前述の i について

$$\sum_j a_{ij} x_j = \lambda x_i$$

と書くことができる。また、上式の対角成分を移行することで

$$\sum_{j \neq i} a_{ij} x_j = \lambda x_i - a_{ii} x_i$$

と変形することができる。x_i は非ゼロとしたため、両辺を x_i で割り絶対値をとることで

$$|\lambda - a_{ii}| = \left| \frac{\sum_{j \neq i} a_{ij} x_j}{x_i} \right| \leq \sum_{j \neq i} \left| \frac{\sum_{j \neq i} a_{ij} x_j}{x_i} \right| \leq \sum_{j \neq i} \left| \sum_{j \neq i} a_{ij} \right| = R_i$$

となり、定理が示された。ここで、三角不等式と $\left|\frac{x_j}{x_i}\right| \leq 1$ を用いた。 □

定理 2 : ある遷移行列 P の任意の右固有値の絶対値は 1 以下である。

Proof. ゲルシュゴリンの定理から任意の右固有値に対してある i が存在し

$$|\lambda - p_{ii}| \leq \sum_{j \neq i} \left| p_{ij} \right|$$

である。P が遷移行列であることから

$$\sum_{j \neq i} \left| p_{ij} \right| \leq 1$$

であり、$\sum_{j \neq i} \left| p_{ij} \right| = 1$ であるとき $p_{ii} = 0$ となるため示される。

また $\sum_{j \neq i} \left| p_{ij} \right| < 1$ であるならば $\left| \lambda - p_{ij} \right| < 1$ であり、

$$|\lambda - p_{ii}| < \sum_{j \neq i} \left| p_{ij} \right| \Leftrightarrow \begin{cases} -\lambda + p_{ii} < \sum_{j \neq i} \left| p_{ij} \right|, & \lambda < p_{ii} \\ \lambda - p_{ii} < \sum_{j \neq i} \left| p_{ij} \right|, & \text{otherwise} \end{cases}$$

であるため示された。 □

定理 3 : ある遷移行列 P は必ず左固有値 1 を持つ。

Proof. P が遷移行列であることがら、右固有値 1 とそれに対する固有ベクトル $\mathbf{1}$ が存在する。また、正方行列の右固有値と左固有値は一致するため左固有値 1 が存在する。 □

系 1 : 遷移行列 P に対して定常分布は必ず存在する。

<u>定理 **4：**</u> ある遷移行列 P が既約かつ非周期であれば

$$\lim_{k\to\infty} P^k = \mathbf{1}\boldsymbol{\pi}$$

となる。ここで $\mathbf{1}$ は各成分が 1 である列ベクトルとした。

Proof. □

以下に定常状態となる例を示す。シミュレーテッド・アニーリングにおいて、ある状態 i で状態 j が選ばれたときに j へと遷移する確率 p'_{ij} は

$$p'_{ij} = \begin{cases} 1, & f(j) \le f(i) \\ \exp\left(-\frac{f(j)-f(i)}{T}\right), & f(j) > f(i) \end{cases}$$

ここで、T はシミュレーテッド・アニーリングで用いられるパラメータであり、f は最小化する目的関数である。このとき、T が不変であるとし、状態 i のときに状態 j が選ばれる確率が一様であるとすれば、このシミュレーテッド・アニーリングの状態遷移は斉次的なマルコフ連鎖となり、遷移行列 $P = \{p_{ij}\}$ は

$$p_{ij} = \begin{cases} 1 - \sum_{j\neq i} p_{ij}, & i = j \\ \frac{1}{n}, & i \neq j, f(j) \le f(i) \\ \frac{1}{n}\exp\left(-\frac{f(j)-f(i)}{T}\right), & i \neq j, f(j) > f(i) \end{cases}$$

と書くことができる。また P は正方行列であり左固有ベクトル 1 が存在するため、定常分布が存在する。

この遷移行列 P によって表わされる定常分布 π がボルツマン分布

$$\pi_i = \frac{\exp\left(-\frac{f(i)}{T}\right)}{\sum_j \exp\left(-\frac{f(j)}{T}\right)}$$

となることを示す。ベクトル π に右から遷移行列を作用させると

$$\pi'_j = \sum_i \pi_i p_{ij}$$

となる。この式を変形すると

$$
\begin{aligned}
\pi'_j &= \sum_i \pi_i p_{ij} \\
&= \pi_j p_{jj} + \sum_{i \neq j} \pi_i p_{ij} \\
&= \pi_j \left(1 - \sum_{k \neq j} p_{jk} \right) + \sum_{i \neq j} \pi_i p_{ij} \\
&= \pi_j - \pi_j \sum_{k \neq j} p_{jk} + \sum_{i \neq j} \pi_i p_{ij} \\
&= \pi_j - \sum_{i \neq j} \pi_j p_{ji} + \sum_{i \neq j} \pi_i p_{ij} \\
&= \pi_j + \sum_{i \neq j} \left(\pi_i p_{ij} - \pi_j p_{ji} \right) \\
&= \pi_j + \sum_{i \neq j} \left(p_{ij} \frac{\exp\left(-\frac{f(i)}{T}\right)}{\sum_k \exp\left(-\frac{f(j)}{T}\right)} - p_{ji} \frac{\exp\left(-\frac{f(j)}{T}\right)}{\sum_k \exp\left(-\frac{f(j)}{T}\right)} \right) \\
&= \pi_j + \frac{1}{\sum_k \exp\left(-\frac{f(j)}{T}\right)} \sum_{i \neq j} \left(p_{ij} \exp\left(-\frac{f(i)}{T}\right) - p_{ji} \exp\left(-\frac{f(j)}{T}\right) \right)
\end{aligned}
$$

となる。$f(i) = f(j)$ のときは $p_{ij} = p_{ji} = \frac{1}{n}$ であるため、$\pi'_j = \pi_j$ となる。$f(j) < f(i)$ のときは

$$
\begin{aligned}
p_{ij} \exp\left(-\frac{f(i)}{T}\right) - p_{ji} \exp\left(-\frac{f(j)}{T}\right) &= \frac{1}{n} \exp\left(-\frac{f(j) - f(i)}{T}\right) \exp\left(-\frac{f(i)}{T}\right) - \frac{1}{n} \exp\left(-\frac{f(j)}{T}\right) \\
&= \frac{1}{n} \exp\left(-\frac{f(j)}{T}\right) - \frac{1}{n} \exp\left(-\frac{f(j)}{T}\right) \\
&= 0
\end{aligned}
$$

となり、$f(j) > f(i)$ のときも同様の変形が行えるため $\pi'_j = \pi_j$ となる。よって遷移行列 P の定常状態としてボルツマン分布 π が存在することが示された。

参考文献

[1] “マルコフ連鎖-Wikipedia”, https://ja.wikipedia.org/wiki/マルコフ連鎖, 参照 May.16,2020.

[2] “確 率 行 列-Wikipedia”, https://ja.wikipedia.org/wiki/確 率 行 列, 参 照 May.16,2020.

[3] “ゲルシュゴリンの定理-Wikipedia”，https://ja.wikipedia.org/wiki/ゲルシュゴリンの定理，参照 May.16,2020.