Short notes on Markov state model

永井哲郎 *

2018 年 11 月 23 日

1　確率過程

離散時間上で定義された確率変数の列として X_t, $t = 0, \tau, 2\tau, \ldots$ を考える．ここで τ は適当な時間間隔であり，素事象が生じる間隔や観測の間隔に対応する．例えば原点 0 から始まる歩幅 1 の時間間隔 τ で生じる酔歩はこの枠組みで書き表せる．

$$X_0 = 0, X_\tau = \pm 1, X_{2\tau} = \pm 2, 0, X_{3\tau} = \pm 3, \pm 1, \ldots, \tag{1}$$

あるいはタンパク質の 2 状態を考える．折りたたみ状態を 1，変性状態を 2 として時間間隔 τ で観測を行えば

$$X_0 = 2, X_\tau = 2, X_{2\tau} = 1, X_{3\tau} = 1, \ldots, \tag{2}$$

と言った観測値の列を得るが，これも離散時間上で定義された確率変数の列として記述できている．

このような確率変数の系列を確率過程と呼ぶ．ここでは起きうる事象の数は有限可算個としておく．つまり，X_t はどれだけ長くても良いのだが，X_t のとりえる値は N 個であるとする．このとりえる値の集合を $\{1, 2, 3, \ldots, N\}$ としておく．将来の状態が過去によらずに現在の状態だけで決まる性質をマルコフ性と呼び，このような性質を満たす確率過程のうち，起きうる事象の数が有限個のものを（離散時間）マルコフ鎖と呼ぶ．このノートでは，この（離散時間）マルコフ鎖を考える．

* 名古屋大学大学院理学研究科

1.1　マルコフ性

（離散時間）マルコフ鎖では，$m, n > 0$ に対して

$$P(X_{(n+m)\tau} = j | X_{n\tau} = i, X_{(n-1)\tau} = x_{n-1}, \cdots, X_{0\tau} = x_0) = P(X_{(n+m)\tau} = j | X_{n\tau} = i) \tag{3}$$

が成立する．つまりある時点より未来の確率はその時点のみの状態で決まりその時点より過去の情報は必要ない．ここで $P(A|B) = P(A \cap B)/P(B)$ は条件付き確率を表す．

1.2　時間に対して一様

時間に対して一様であることも要請しておく．

$$P(X_{(n+m)\tau} = j | X_{n\tau} = i) = P(X_{m\tau} = j | X_0 = i) \tag{4}$$

$$\equiv P(i, j; m\tau) \tag{5}$$

特に

$$P(i, j; \tau) \equiv p(i, j) \tag{6}$$

と書くことにする．

1.3　遷移行列

$p(i, j)$ を並べて作った行列を遷移行列と呼ぶ．

$$T_{ij} \equiv p(i, j) \tag{7}$$

遷移行列は明らかに $T_{ij} \geq 0$ と $\sum_j T_{ij} = 1$ とを満たす．（逆に実正方行列 T_{ij} がこの二つを満たせば T_{ij} は遷移行列であり対応するマルコフ鎖が作れる．）

2　性質

2.1　Chapman-Kolmogorov 等式

マルコフ性より Chapman-Kolmogorov 等式が成り立つ：

$$P(i, j; (m+n)\tau) = \sum_k P(i, k, m\tau) P(k, j, n\tau) \tag{8}$$

2.1.1 系

時間 2τ 進む時,

$$P(X_{2\tau}=j|X_0=i)=\sum_k P(X_{2\tau}=j|X_\tau=k)P(X_\tau=k|X_0=i) \tag{9}$$

$$=\sum_k p(i,k)p(k,j) \tag{10}$$

$$=(T^2)_{ij} \tag{11}$$

一般に $m\tau$ 進む時,

$$P(X_{m\tau}=j|X_0=i)=\sum_k P(i,k;(m-1)\tau)P(k,j;\tau) \tag{12}$$

$$=\sum_k P(i,k;(m-1)\tau)p(k,j) \tag{13}$$

となる．以上から直ちに,

$$P(X_{m\tau}=j|X_0=i)=(T^m)_{ij} \tag{14}$$

2.2 定常分布

もしも時間発展に対して不変な確率分布

$$\sum_i \mu(i)p(i,j)=\mu(j) \tag{15}$$

が存在するとき，これを定常分布 $\boldsymbol{\mu}$ と呼ぶ．（このノートでは $\boldsymbol{\mu}$ は横ベクトル．）

2.3 収束定理

適当な条件のもと（T が既約，非周期的かつ定常分布を持つとき），T^n は μ に収束する．

$$\lim_{n\to\infty}(T^n)_{ij}=\mu(j) \tag{16}$$

2.3.1 例

遷移行列

$$T = \begin{pmatrix} 0.75 & 0.15 & 0.1 \\ 0.25 & 0.65 & 0.1 \\ 0.25 & 0.15 & 0.6 \end{pmatrix} \tag{17}$$

に対して定常分布 $\boldsymbol{\mu} = (0.5, 0.3, 0.2)$ が存在するが

$$T^8 = \begin{pmatrix} 0.50195312 & 0.29882812 & 0.19921875 \\ 0.49804688 & 0.30273438 & 0.19921875 \\ 0.49804688 & 0.29882812 & 0.203125 \end{pmatrix} \tag{18}$$

さらに

$$T^{16} = \begin{pmatrix} 0.50000763 & 0.29999542 & 0.19999695 \\ 0.49999237 & 0.30001068 & 0.19999695 \\ 0.49999237 & 0.29999542 & 0.20001221 \end{pmatrix} \tag{19}$$

となり，行列 T^n の ij 成分は j のみで決まる定常分布に近づく様子が分かる．

2.4 平均に対する定理

適当な条件のもと（T が既約かつ定常分布を持つとき），

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{m=n} f(X_m) = \sum_i \mu(i) f(i) \tag{20}$$

が満たされ，時間平均と定常状態に基づく平均は一致する．

2.5 詳細釣り合い条件と可逆性

次の条件を詳細釣り合いの条件（detailed balance condition）と呼ぶ．

$$\mu(i)p(i,j) = \mu(j)p(j,i) \tag{21}$$

詳細釣合いを満たすときマルコフ鎖は可逆である．つまり，$Y_m = X_{n-m}$ と構成した確率過程は，X_m と同じ遷移確率 $p(i,j)$ で発展させた過程となる．

2.5.1 可逆でない例

次の遷移行列は，定常分布を持つが可逆ではない例である．

$$T = \begin{pmatrix} 0.1 & 0.9 & 0.0 \\ 0.0 & 0.1 & 0.9 \\ 0.9 & 0.0 & 0.1 \end{pmatrix} \tag{22}$$

実際に $\boldsymbol{\mu} = (1/3, 1/3, 1/3)$ は定常分布となる．しかし詳細釣り合い条件を満たさないので可逆ではない．可逆ではないとき，状態空間の中に 1 方向に確率が流れる輪が形成される．この例では状態 1 から 2 へ 2 から 3 へ 3 から 1 への循環がある．そのため詳細釣り合い条件を満たさない遷移行列で表現されるマルコフ鎖は平衡状態のモデルとしては都合が悪い．そのため，対応する平衡状態を持つマルコフ鎖であれば詳細釣り合い条件は満たされているべきである．

3 固有値問題

定常状態は $\boldsymbol{\mu}$ は $\boldsymbol{\mu}T = \boldsymbol{\mu}$ で決まるので，$\boldsymbol{\mu}$ は固有値問題

$$|\lambda I - T| = 0 \tag{23}$$

の $\lambda = 1$ に対する固有ベクトルとして与えられる．

特に詳細釣り合い条件を満たすときは対称行列の対角化を用いて固有値・固有ベクトルが計算可能である．詳細釣り合い条件より，

$$\mu(i)T_{ij} = \mu(j)T_{ji} \tag{24}$$

$$\mu(i)^{1/2}T_{ij}\mu(j)^{-1/2} = \mu(j)^{1/2}T_{ji}\mu(i)^{-1/2} \tag{25}$$

より，$D_{ij} = \delta_{ij}\mu(i)^{1/2}$ を用いて定義される DTD^{-1} は実対称行列であり実数の固有値を持ちかつ対角化可能．DTD^{-1} の固有値と固有ベクトルを $\lambda_1, \lambda_2, \ldots$ と $\boldsymbol{a}_1, \boldsymbol{a}_2 \ldots$ として

$$\boldsymbol{a}_i DTD^{-1} = \lambda_i \boldsymbol{a}_i \tag{26}$$

$$\boldsymbol{a}_i DT = \lambda_i \boldsymbol{a}_i D \tag{27}$$

を得て，T の固有値は，λ_i，対応する固有ベクトルは $\boldsymbol{x}_i = \boldsymbol{a}_i D$ となる．

詳細は立ち入らないがここで適当な条件のもとペロン・フロベニウスの定理を用いると $1 = \lambda_1 > \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_N > -1$ となり，λ_1 の重複度は 1 で他の固有値の絶対値は真

に 1 よりも小さいことが示される．また，$\boldsymbol{x}_1(=\boldsymbol{\mu})$ の各成分は全て正であることも示される．また，他の固有ベクトルは

$$\boldsymbol{x}_\ell T = \lambda \boldsymbol{x}_\ell \tag{28}$$

$$\sum_i x_{\ell i} T_{ij} = \lambda x_{\ell j} \tag{29}$$

$$\sum_{i,j} x_{\ell i} T_{ij} = \lambda \sum_j x_{\ell j} \tag{30}$$

$$\sum_i x_{\ell i} = \lambda \sum_j x_{\ell j} \tag{31}$$

より $\lambda \neq 1$ であることから $\sum_i x_{\ell i} = 0$ となる．このため $\boldsymbol{x}_j$ $(j > 1)$ は平衡状態からのズレを表し，状態遷移のモードと見なせる．実際に固有ベクトルを基底として任意の状態 $\boldsymbol{\xi}$ は

$$\boldsymbol{\xi}(t=0) = \boldsymbol{\mu} + \sum_{i=2}^{N} c_i \boldsymbol{x}_i \tag{32}$$

$$\boldsymbol{\xi}(t=\tau) = \boldsymbol{\xi} T = \boldsymbol{\mu} + \sum_i c_i \lambda \boldsymbol{x}_i \tag{33}$$

$$\boldsymbol{\xi}(t=n\tau) = \boldsymbol{\xi} T^n = \boldsymbol{\mu} + \sum_i c_i \lambda^n \boldsymbol{x}_i \tag{34}$$

$\lambda > 0$ の時 $\lambda_i = \exp(-\tau/\tau_i)$ として

$$\boldsymbol{\xi}(t=n\tau) = \boldsymbol{\mu} + \sum_i c_i \exp(-n\tau/\tau_i) \boldsymbol{x}_i \tag{35}$$

τ_i は “implied timescale” と呼ばれる．λ_i が 1 に近い (τ_i が大きい) ほどモード i に対する緩和は遅い．

また，マルコフ過程に対して $T(n\tau) = T^n$ が成立するので $T(n\tau)$ の固有値 λ'_i は λ^n と一致する．そのため，$-n\tau/\log\lambda'_i = -\tau/\log\lambda_i$ となり，マルコフ鎖であれば $-n\tau/\log\lambda'_i$ は n に対して不変となる（逆は成立しないが Markov 性の検証に使われる）．

4　マルコフモデルの作り方

まず，シミュレーションなどで得られた情報を元に状態を離散化する．離散化の方法は成書を参考のこと．シミュレーションや実験結果から T_{ij} を求める素朴な方法は，$T_{ij} = P(X_{t+\tau} = j | X_t = i) = P(X_{t+\tau} = j, X_t = i)/P(X_t = i)$ なので，時間間隔 τ で実際に起きた状態 i から状態 j への遷移の数を状態 i が発生した数で割ればよい．より進ん

だ遷移行列の計算方法は成書を参考のこと．時間間隔 τ の決め方も成書を参考のこと．(元々 Markov 性の成り立つ系でも粗視化された状態に対するダイナミクスは τ が短いと Markov 性が一般には担保されない．) T_{ij} が求まれば T_{ij} を対角化することで，平衡分布や遷移モード，遷移モードの implied timescale などが求まる．得られたモードの解析方法についても成書を適切に参照すること．

参考文献

[1] 線形代数入門，藤原正彦，東京大学出版会．

[2] 線形代数，長谷川浩司，日本評論社．

[3] 計算統計 II，伊庭幸人ほか，岩波書店．

[4] "An Introduction to Markov State Models and their Application to Long Timescale Molecular Simulation," G.R. Bowman et al., Springer.

[5] "Essentials of Stochastic Processes," R. Durrett, Springer.