原子力研究開発機構夏季実習生 repor

2023 年 10 月 12 日

目次

第 I 部
確率過程

1 Introduction

ここでは Markov 過程について考える．Markov 過程は主に次の 2 種類に区別される．

定義 1.1（マルコフ連鎖 (Markov chain)） 離散時間における離散状態を取る Markov 過程のこと

定義 1.2（マルコフジャンプ過程 (Markov jump process)） 連続時間における離散状態を取る Markov 過程のこと

2 Markov process with discrete state

2.1 Markov chain

■基本的な定義 とびとびの状態を $j = 1, 2, \ldots, \Omega$ とおく．ある時点で状態が j である確率を p_j とする．確率 p_j は規格化条件

$$0 \le p_j \le 1, \quad \sum_{j=1}^{\Omega} p_j = 1 \tag{1}$$

を満たす．確率 p_j を並べると，確率分布 (確率ベクトル) は,

$$\boldsymbol{p} = \begin{pmatrix} p_1 \\ p_2 \\ \vdots \\ p_\Omega \end{pmatrix} = (p_j)_{j=1,\ldots,\Omega} \tag{2}$$

と書くことができる．状態が k から j に移る確率 $T_{j,k}$ を遷移確率，推移確率 (transition probability) と呼ぶ．そして，遷移確率を成分とする $\Omega \times \Omega$ 行列

$$T = (T_{j,k})_{j,k=1,\ldots,\Omega} \tag{3}$$

を確率行列と呼ぶ．確率行列の成分である，遷移確率は

$$0 \leq T_{j,k} \leq 1 \tag{4}$$

と，任意の k に対して，規格化条件

$$\sum_{j=1}^{\Omega} T_{j,k} = 1 \tag{5}$$

を満たす．

■確率の保存則　任意の確率行列 T と任意の Ω この成分をもつ列ベクトル $\boldsymbol{v}$ について，

$$\sum_{j=1}^{\Omega} (T\boldsymbol{v})_i = \sum_{j,k=1}^{\Omega} T_{j,k}(\boldsymbol{v})_k = \sum_{j=1} (\boldsymbol{v})_j \tag{6}$$

が成り立つ．ここで，2 番目の等式で，規格化条件 (5) を用いた．

3　マルコフ連鎖 Markov chain

離散時間 $n = 0, 1, 2, \ldots$ における確率行列 $T^{(n)}$，$n = 1, 2, \ldots$ を考える．
基本的なアイデア：時刻 $n-1$ に系が状態 k にいたとすると，そのとき，時刻 n に系が状態 j にいる確率行列は，$T_{j,k}^{(n)}$ となる．
確率過程：状態 j は確率的に決定される．
マルコフ過程：直前の時刻の状態が次の状態を決める．

時刻 n に系が状態 j にいる確率を $p_j^{(n)}$ とする．時刻 n での確率分布を $\boldsymbol{p}^{(n)} = (p_j^{(n)})_{j=1,\ldots,\Omega}$ とする．時刻 $n-1$ に系が状態 k にいる確率に確率行列 $T_{j,k}^{(n)}$ をかければ，時刻 n に系が状態 j にいる確率が得られる：

$$p_j^{(n)} = \sum_{k=1}^{\Omega} T_{j,k}^{(n)} p_k^{(n-1)} \tag{7}$$

確率分布で書くと，

$$\boldsymbol{p}^{(n)} = T^{(n)} \boldsymbol{p}^{(n-1)} \tag{8}$$

である．つまり，時間発展の規則は，確率行列をかければよい．系が初期状態の場合の確率分布を $\boldsymbol{p}^{(0)} = (p_1^{(0)}, p_2^{(0)}, \ldots, p_\Omega^{(0)})^t$ と記述する．すると，時刻 1 での確率分布は

$\boldsymbol{p}^{(1)} = T^1\boldsymbol{p}^{(0)}$ であり，以下，$\boldsymbol{p}^{(2)} = T^2\boldsymbol{p}^{(1)}$ のように，次々と次の時刻での確率分布が決定される．したがって，

$$\boldsymbol{p}^{(n)} = T^{(n)}T^{(n-1)}\cdots T^{(1)}\boldsymbol{p}^{(0} \tag{9}$$

と書くことができる．k のように確率分布が時間と共に確率的に変化するのが，有限離散状態・離散時間のマルコフ連鎖である．

■単調性　初期分布は異なるが，状態を遷移させるために用いる確率行列は等しいとする．このとき，

$$\boldsymbol{p}^{(n)} = T^{(n)}\boldsymbol{p}^{(n-1)},\ \boldsymbol{q}^{(n)} = T^{(n)}\boldsymbol{q}^{(n-1)} \tag{10}$$

このとき，相対エントロピーは

$$D(\boldsymbol{p}|\boldsymbol{q}) \geq D(T\boldsymbol{p}|T\boldsymbol{q}) \tag{11}$$

が成り立つので，

$$D(\boldsymbol{p}^{(n-1)}|\boldsymbol{q}^{(n-1)}) \geq D(\boldsymbol{p}^{(n)}|\boldsymbol{q}^{(n)}) \tag{12}$$

が成り立つ．つまり，2 つの確率分布の距離は決して増えることはないということがわかる．

■経路の確率 path prob　初期分布が $\boldsymbol{p}^{(0)}$ であり，マルコフ連鎖によって時刻 n まで時間発展したとき，時刻 $0, 1, \ldots, n$ において，系が状態 $j_0, j_1, \ldots, j_n$ である確率は，

$$p_{j_0,j_1,\ldots,j_n} \equiv T^{(n)}_{j_n,j_{n-1}}T^{(n-1)}_{j_{n-1},j_{n-2}}\cdots T^{(1)}_{j_1,j_0}p^{(0)}_{j_0} \tag{13}$$

と定義される．これは，図に示すように，歴史の確率，経路の確率を表しているといえる．そして，この確率は，

$$\sum_{j_0,\ldots,j_n=1}^{\Omega} p_{j_0,j_1,\ldots,j_n} = 1 \tag{14}$$

を満たし，確率分布で書けば，

$$\boldsymbol{p} = (p_{j_0,\ldots,j_n})_{j_0,\ldots,j_n=1,\ldots,\Omega} \tag{15}$$

$\{1, \ldots, \Omega\}^n$

3.1 詳細釣り合いの条件 (detailed balance condition)

まず目的分布として，定常分布が与えられるとする．この分布をを実現するような遷移確率は何かを考えるという状況を考える．定常分布 $\boldsymbol{p}^{(s)} = (p_j^{(s)})_{j=1,\ldots,\Omega}$ が与えられたとする．すべての $j = 1,\ldots,\Omega$ に対して，$p_j^{(s)} > 0$ とする．ここで，詳細つり合いの条件を導入する：

定義 3.1（詳細釣り合いの条件 (detailed balance condition)） もしも，任意の $j \neq k$ に対して

$$T_{j,k}\boldsymbol{p}_k^{(s)} = T_{k,j}\boldsymbol{p}_j^{(s)} \tag{16}$$

を満たすような確率分布 $\boldsymbol{p}_j^{(s)}$ が存在するとする．このとき，遷移確率 $T_{j,k}$ $j,k = 1,\ldots,\Omega$ は $\boldsymbol{p}_j^{(s)}$ について詳細釣り合い条件を満たすという．

詳細釣り合い条件を満たしているとき，

$$\begin{aligned}\sum_{k=1}^{\Omega} T_{j,k}p_k^{(s)} &= T_{j,j}p_k^{(s)} + \sum_{k(\neq j)} T_{j,k}p_k^{(s)} \\ &= T_{j,j}p_k^{(s)} + \sum_{k(\neq j)} T_{k,j}p_j^{(s)} \\ &= \left(\sum_{k}^{\Omega} T_{k,j}\right)p_j^{(s)} = p_j^{(s)} \end{aligned} \tag{17}$$

が成り立つ．ここで，2 つ目の等号で詳細釣り合い条件を使った．すなわち，

$$T\boldsymbol{p}^{(s)} = \boldsymbol{p}^{(s)} \tag{18}$$

を得る．逆に $T\boldsymbol{p}^{(s)} = \boldsymbol{p}^{(s)}$ であるから詳細釣り合い条件が成り立つとは限らない．つまり，詳細釣り合い条件は $\boldsymbol{p}^{(s)}$ が定常分布になるための十分条件であることがわかる．一般に，

$$T\boldsymbol{p}^{(s)} = \boldsymbol{p}^{(s)} \tag{19}$$

$$\sum_{k=1}^{\Omega} T_k p_k^{(s)} = p_k^{(s)} \tag{20}$$

をつり合い条件 (balanced condition) と呼ぶ．定常分布を用意するための，T の決め方はいくらでも考えうることができるのだが，その中の一つ (one of them) が詳細つり合い条件であるということに注意が必要である．詳細つり合い条件はとてもシンプルで扱いやすいアイデアなのだが，これを満たすモンテカルロ法は非常に遅いというのが弱点がある．また，詳細つり合い条件は T の決め方の一つであったから，それ以外にもつり合い条件満たし，T を決定する方法は存在する．つまり，詳細つり合い条件を破りながら，つり合い条件を満たす手法もあるということに注意が必要である．次では，目的の分布を用意するための数値計算手法であるマルコフ連鎖モンテカルロ法について解説を行う．また，具体的な実装法についても述べる．

3.2 マルコフ連鎖モンテカルロ法 (Markov chain monte calro method, MCMC)

Example: ボルツマン分布

例として，ボルツマン分布の場合を考える．ボルツマン分布を目標分布に設定する：

$$P(\vec{x}) \equiv q_\theta(\vec{x}) = \frac{1}{Z} \exp\left(-E_\theta(\vec{x})\right) \tag{21}$$

ボルツマン分布に従い，遷移確率の比を計算する：

$$\frac{w(\vec{x}|\vec{x}')}{w(\vec{x}'|\vec{x})} = \frac{P(\vec{x})}{P(\vec{x}')} = \frac{\exp\left(-E_\theta(\vec{x})\right)}{\exp\left(-E_\theta(\vec{x}')\right)} = \exp\left\{-(E_\theta(\vec{x}) - E_\theta(\vec{x}'))\right\} \tag{22}$$

ここで，左辺の意味について説明する．

$$\exp\left\{-(\overbrace{\underbrace{E_\theta(\vec{x})}_{\text{フリップ後}} - \underbrace{E_\theta(\vec{x}')}_{\text{フリップ前}}}^{\text{熱}:\Delta E(\vec{x}|\vec{x}')})\right\} \tag{23}$$

ここでは，$\vec{x}'$, $\vec{x}$ はそれぞれ，フリップ前，フリップ後のスピンの取り得る状態を表し，物質の内部的な変化に対応している．つまり，θ を固定して，スピンの状態を変化させた後のエネルギーの変化 $\Delta E(\vec{x}|\vec{x}')$ は熱力学とのアナロジーから熱に対応していることがわかる．一方，θ は人が外から加え変化さえる外部パラメータである（例えば，外場や相互作用）．つまり，内部状態は変えずに，外部パラメータ θ を動かした後のエネルギー変化 $E_{\theta'} - E_\theta$ は仕事に対応している．

ここで，遷移確率 $w(\vec{x}|\vec{x}')$ を提案 (propose) 確率 $C(\vec{x}|\vec{x}')$ と受理 (accept) 確率 $A(\vec{x}|\vec{x}')$ に分離する：遷移確率 = 提案確率 × 受理確率

$$w(\vec{x}|\vec{x}') = C(\vec{x}|\vec{x}') \times A(\vec{x}|\vec{x}'). \tag{24}$$

ここで，提案確率 $C(\vec{x}|\vec{x}')$ は自分でルールを課して自由に決めることができる確率のことである．例えば，提案確率はスピンの状態を $\vec{x}' \to \vec{x}$ に変化させるルールを提案する．そして，受理確率はその変化を受け入れるかに OK or NG を出す役目を果たす．

例：1spin flip rule

　このとき，N 個のスピンのどれか一つをフリップさせることを提案する．つまり，提案確率は

$$C(\vec{x}|\vec{x}') = 1/N \tag{25}$$

である．例えば，3 つのスピン系を考える．2 番目のスピンを -1 から $+1$ に変えることを提案する：

$$\vec{x}' = \{+1,\ -1,\ +1\} \qquad \to \qquad \vec{x} = \{+1,\ +1,\ +1\} \tag{26}$$

■メトロポリス法 (metropolis method)　次に，受理確率の決め方の一つである，メトロポリス法について説明する．

定義 3.2（メトロポリス法）　メトロポリス法において受理確率 $A(\vec{x}|\vec{x}')$ は次のように定義される：

$$A(\vec{x}|\vec{x}') = \min\left\{1,\ \frac{P(\vec{x})}{P(\vec{x}')}\frac{C(\vec{x}'|\vec{x})}{C(\vec{x}|\vec{x}')}\right\} \tag{27}$$

ここで，

$$\frac{P(\vec{x})}{P(\vec{x}')}\frac{C(\vec{x}'|\vec{x})}{C(\vec{x}|\vec{x}')} < 1 \tag{28}$$

である．

メトロポリス法は詳細つり合い条件を満たすための手法の 1 つである．そこで，メトロポリス法が詳細つり合い条件を満たすかどうかを確認する．まず，遷移確率 $w(\vec{x}|\vec{x}')$ を計算する：

$$w(\vec{x}|\vec{x}') = C(\vec{x}|\vec{x}') \times \frac{P(\vec{x})}{P(\vec{x}')}\frac{C(\vec{x}'|\vec{x})}{C(\vec{x}|\vec{x}')} \tag{29}$$

逆の遷移確率は

$$w(\vec{x}'|\vec{x}) = C(\vec{x}'|\vec{x}) \times 1 \tag{30}$$

となる．ここで，逆の受理確率

$$A(\vec{x}'|\vec{x}) = \frac{P(\vec{x}')}{P(\vec{x})}\frac{C(\vec{x}|\vec{x}')}{C(\vec{x}'|\vec{x})} > 1 \tag{31}$$

がメトロポリス法の受理確率が 1 を超えてしまうため，メトロポリス法の受理確率の定義から 1 に取った．よって

$$\frac{w(\vec{x}|\vec{x}')}{w(\vec{x}'|\vec{x})} = \frac{C(\vec{x}|\vec{x}') \times \dfrac{P(\vec{x})}{P(\vec{x}')}\dfrac{C(\vec{x}'|\vec{x})}{C(\vec{x}|\vec{x}')}}{C(\vec{x}'|\vec{x}) \times 1} = \frac{P(\vec{x})}{P(\vec{x}')} \tag{32}$$

したがって，詳細つり合い条件を満たすことがわかる．ただ，何度も言うが，詳細つり合い条件を満たしている必要はないということである．

3.3 具体的な実装方法

N 個のスピンから 1 つスピンを選びフリップさせるルールを考える．このとき提案確率は $C(\vec{x}|\vec{x}') = 1/N$ となる．まず，一様分布に従う乱数 $r \in [0,1)$ を発生させる．そして，1 スピンフリップさせた後のエネルギーの変化 $\Delta E(\vec{x}|\vec{x}')$ を調べる．ここから，受理確率を実現することができる．まずエネルギーが増える場合，$\Delta E(\vec{x}|\vec{x}') > 0$ の場合，乱数 r が受理確率 $P(\vec{x})P(\vec{x}')$ よりも小さくなれば，1 スピンフリップ $\vec{x}' \to \vec{x}$ を受け入れ (accept)，それ以外ならば，棄却 (reject) する：

$$\Delta E(\vec{x}|\vec{x}') > 0 \quad \begin{cases} r \leq \dfrac{P(\vec{x})}{P(\vec{x}')} = \exp\Big\{-(E_\theta(\vec{x}) - E_\theta(\vec{x}'))\Big\}, & \text{なら accepe} \\ \text{otherwise}, & \text{なら reject} \end{cases} \tag{33}$$

一方で，エネルギー変化が負の場合，$\Delta E(\vec{x}|\vec{x}') < 0$ の場合，受理確率はメトロポリス法の定義より 1 を超えてしまう：

$$A(\vec{x}'|\vec{x}) = \frac{P(\vec{x}')}{P(\vec{x})} > 1 \tag{34}$$

したがって，受理確率は最小の 1 となり，つまり 100% accept される．

マルコフ連鎖モンテカルロ法は勾配降下法の拡張になっている．もしも，$\Delta E(\vec{x}|\vec{x}') < 0$ のみを許せば，それは勾配降下法に帰着する．勾配降下法では，エネルギーを下げる方向の変化のみを許していた．しかし，MCMC では，$\Delta E(\vec{x}|\vec{x}') > 0$ の場合，すなわち，エネルギーが高くなる方向も確率的に $e^{-\Delta E(\vec{x}|\vec{x}')}$ で許しているということを表している．少し低い山ならば登ることを許しているということである．

4 Markov jump process

この節では，連続時間における離散状態に関する Markov 過程，すなわち Markov ジャンプ過程について考える．

4.1 定義

系は離散状態 $j = 1, \ldots, \Omega$ を取る．系が時刻 t に状態 j を取る確率を $p_j(t)$ とおく．[*1] 確率 $p_j(t)$ は任意の時刻 t で規格化条件を満たすとする

$$\sum_{j=1}^{\Omega} p_j(t) = 1. \tag{35}$$

時間 t は連続に流れていき，系の状態は，ある瞬間に，ある状態から別の状態へと一瞬でジャンプするとする．このようなジャンプのおこる割合は，過去の記憶に影響されず (Markov 性)，その瞬間の系の状態だけで決まるとする．

ある時刻に系が状態 j にいる場合を考える．ここである状態，遷移率 (transition probability) と

定義 4.1（遷移率 (transition rate)） ある時刻に系が状態 i にいるとする．それから短い時間間隔 Δt の間に系が別の状態 i へ遷移している確率を次のように定める．

$$\Delta t \; \omega_{i \to j} + \mathcal{O}((\Delta t)^2) \tag{36}$$

このとき，単位時間あたりに状態 i から j へ遷移する割合を $\omega_{i \to j}$ と書き，これを遷移率 (transition rate) と呼ぶ．

定義 4.2（escaoe rate） ある時刻に系が状態 j にいるとする．このとき，状態 j か

[*1] 離散時間の場合は数列として，連続時間の場合は確率分布を時間 t 関数の形として表す．

ら j 以外の別の状態へ逃げていく確率を次のように定義する：

$$\lambda_j(t) = \sum_{k(\neq k)} \omega_{j\to k} \geq 0 \tag{37}$$

これを escape rate と呼ぶ.

4.2 Master equation

状態 j の $t \sim t+\Delta t$ の時間発展を考える．時刻 $t+\Delta t$ に状態 j にいる確率は次のように記述される：

$$p_j(t+\Delta t) = -\left\{\Delta\ \lambda_j(t) + \mathcal{O}((\Delta t)^2)\right\} p_j(t) + \sum_{k(\neq j)} \left\{\Delta\ \omega_{k\to j}(t) + \mathcal{O}((\Delta t)^2)\right\} p_k(t) + p_j(t) \tag{38}$$

右辺第一項 $-\left\{\Delta\ \lambda_j(t) + \mathcal{O}((\Delta t)^2)\right\} p_j(t)$ は状態 j から escape する確率を，第二項 $\sum_{k(\neq j)} \left\{\Delta\ \omega_{k\to j}(t) + \mathcal{O}((\Delta t)^2)\right\} p_k(t)$ は k から j に入ってくる確率を，第三項 $p_j(t)$ は j にそのままとどまっている確率を表す.

この式を次のように変形する：

$$p_j(t+\Delta t) - p_j(t) = -\left\{\Delta\ \lambda_j(t) + \mathcal{O}((\Delta t)^2)\right\} p_j(t) + \sum_{k(\neq j)} \left\{\Delta\ \omega_{k\to j}(t) + \mathcal{O}((\Delta t)^2)\right\} p_k(t) \tag{39}$$

そして両辺を Δt で割り，$\Delta t \to 0$ の極限を取ると，次式を得る：

$$\frac{d}{dt} p_j(t) = -\lambda_j(t) p_j(t) + \sum_{k(\neq j)} \omega_{k\to j}(t) p_k(t). \tag{40}$$

ここで，遷移率行列 (transition rate matrix) を導入する.

定義 4.3（遷移率行列 (transition rate matrix)） 遷移率行列 $R(t) = (R_{j,k})_{j,k=1,\dots,\Omega}$ は遷移率 $\omega_{k\to j}$ と escape rateλ_j を用いて，次のように定義される：

$$R_{j,k}(t) = \omega_{k\to j}(t) \geq 0 \quad (j \neq k) \tag{41}$$

$$R_{k,k}(t) = -\lambda_k(t) \leq 0 \tag{42}$$

また，遷移率行列は任意の k について次を満たす：

$$\sum_{j=1}^{\Omega} R_{j,k} = 0 \tag{43}$$

これは escape rate$\lambda_k(t)$ が $\lambda_k(t) = \sum_{j(\neq k)} \omega_{k\to j}(t)$ と書けるから，$\sum_{j=1}^{\Omega} R_{j,k} = \sum_{j(\neq k)} R_{j,k} + R_{k,k} = 0$ となることからわかる．

遷移率行列 $R(t)$ を用いると，微分方程式 (40) は

$$\frac{d}{dt} p_j(t) = \sum_{k=1}^{\Omega} R_{j,k}\ p_k(t) \tag{44}$$

あるいは,

$$\frac{d}{dt} \boldsymbol{p}(t) = R(t)\boldsymbol{p}(t) \tag{45}$$

と書ける．この式は物理ではマスター方程式 (master equation) と、数学ではコルモゴロフの先進方程式 (Kolmogorov's forward equation) と呼ぶ.

4.3 some basic properties

ここでは Markov ジャンプ過程のいくつかの基本的な性質について述べる．

4.4 Master 方程式の行列表現

4.5 確率流の計算

4.6 convergence theorem for stationary process（定常過程の収束定理）

4.7 メトロポリス法

参考文献