# 確率的感染症モデル　第2回

永幡幸生
新潟大学
nagahata@eng.niigata-u.ac.jp

2022 前期

# 連続時間 Markov 連鎖

離散時間の Markov 連鎖は推移確率行列 $P$ が与えられている時には非常に分かりやすく、コンピュータを使った（モンテカルロ）シミュレーションとの相性が非常に良い。
ここでは離散時間 Markov 連鎖の定義から始めることにする。

# 連続時間 Markov 連鎖

離散時間 Markov 連鎖を定義するには条件付確率が必要になるが、高校の数 I の範囲の割り算で定義するもの、すなわち

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \qquad P(B) \neq 0$$

で与えられるもので十分である。
表記を簡単にするため $P(A, B) = P(A \cap B)$ 等と $\cap$ を省略して書くことにする。

# 連続時間 Markov 連鎖

Markov 連鎖は確率過程の特別なものである。
確率過程とは時間パラメータの入った確率変数であり $(X_n)_{n\geq 0}$ を確率過程とした時 $X_n \in S$ となる $S$ のことを状態空間と呼ぶ。
通常確率過程 $(X_n)_{n\geq 0}$ に対して $X_n$ の確率法則はそれ以前の状態 $X_0, X_1, \dots, X_{n-1}$ に依存する。
(離散時間)Markov 連鎖とは次の Markov 性を満たす確率過程のことである
$P(X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \dots, X_1 = x_1, X_0 = x_0) > 0$ であれば

$$
\begin{aligned}
&P(X_n = x_n, | X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \dots, X_1 = x_1, X_0 = x_0) \\
&= P(X_n = x_n | X_{n-1} = x_{n-1})
\end{aligned}
$$

すなわち $X_n$ の確率法則が 1 歩前の $X_{n-1}$ の状態だけで決定されるものである。

# 連続時間 Markov 連鎖

(離散時間)Markov 過程は $X_n$ の確率法則が 1 歩前の $X_{n-1}$ の状態だけで決定されるものであるが次のように考えるとランダムウォークの自然な拡張と思える。
ランダムウォーカーはコイン (サイコロ) を投げて出た面 (目) によって進む方向を決めて行く。特にこのコイン (サイコロ) はランダムウォーカーの持ち物でありいつも同じものを使っている。
一方で Markov 過程ではウォーカーは自分自身のコイン (サイコロ) をもっておらず、そのたどり着いた場所に備え付けのコイン (サイコロ) を使っている。そのためコイン (サイコロ) 達が場所によって違うことを考える必要がある。

# 連続時間 Markov 連鎖

(人間の考える) 多くの確率過程は (状態空間を大きく取れば) Markov 過程になる。
また、決定論的な運動法則である、ニュートン力学に従うような運動は、現在の位置と、速度が分かれば未来の運動法則が分かる。この位置と速度を状態空間とみなして、対応する確率的な運動法則は現在の状態空間のどこにいるかが分かれば、未来の確率法則が分かる。と読み替えると、これは Markov 連鎖になり、物理現象を中心として多種多様な確率現象をモデル化するのに適している。

# 連続時間 Markov 連鎖

話を単純化するために

$$P(X_{n+1} = y | X_n = x) = P(X_1 = y | X_0 = x), \quad \forall n$$

を仮定する。
この仮定を満たす Markov 連鎖を時間的に一様な Markov 連鎖と呼ぶ。
時間的に一様な Markov 連鎖において

$$P_{x,y} := P(X_1 = y | X_0 = x)$$

とおき (巨大なサイズの) 行列を考える。
この行列を推移確率もしくは推移確率行列と呼ぶ。

# 連続時間 Markov 連鎖

次のような計算をする。

$$
\begin{aligned}
&P(X_{n+m} = y | X_m = x) \\
&= P(X_{n+m} = y, \bigcup_{y_{n+m-1}} X_{n+m-1} = y_{n+m-1} | X_m = x) \\
&= \sum_{y_{n+m-1}} P(X_{n+m} = y, X_{n+m-1} = y_{n+m-1} | X_m = x) \\
&= \sum_{y_{n+m-1}} P(X_{n+m} = y | X_{n+m-1} = y_{n+m-1}) \\
&\qquad \times P(X_{n-1} = y_{n+m-1} | X_m = x) \\
&= \sum_{y_{n+m-1}} P(X_{n+m-1} = y_{n+m-1} | X_m = x) P_{y_{n+m-1}, y}
\end{aligned}
$$

# 連続時間 Markov 連鎖

以下帰納的に計算することで

$$
\begin{aligned}
&P(X_{n+m}=y|X_m=x)\\
&=\sum_{y_{n+m-1}} P(X_{n+m-1}=y_{n+m-1}|X_m=x)P_{y_{n+m-1},y}\\
&\vdots\\
&=\sum_{y_{m+1}}\sum_{y_{m+2}}\cdots\sum_{y_{n+m-1}} P_{x,y_{m+1}}P_{y_{m+1},y_{m+2}}\cdots P_{y_{n+m-1},y}\\
&=(P^n)_{x,y}
\end{aligned}
$$

と行列 $P$ の $n$ 乗に帰着できる。

# 連続時間 Markov 連鎖

このため時間一様な Markov 連鎖においては推移確率 $P$ の $n$ 乗を計算することが重要になるが、これは行列の計算、特に固有値、固有ベクトル、対角化が重要になる。

$$Q^{-1}PQ = \Lambda, \qquad P = Q\Lambda Q^{-1}$$

$$\Lambda = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

と書ければ

# 連続時間 Markov 連鎖

$$P^m = Q\Lambda^m Q^{-1}$$

と書け、

$$\Lambda^m = \begin{pmatrix} \lambda_1^m & & \\ & \ddots & \\ & & \lambda_n^m \end{pmatrix}$$

であるので固有値 $\lambda_i$ の情報が分かれば定性的な性質はある程度簡単に分かる。
簡単のため、このように対角化が可能な推移確率 $P$ が与えられている離散時間 Markov 連鎖をベースにした連続時間 Markov 連鎖を与えることにする。

# 連続時間 Markov 連鎖

抽象的で分かりにくい場合は、一番簡単な例として、状態空間が $S = \{1, 2\}$ と 2 つの要素だけから成り、推移確率行列として

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

ただし $0 < \alpha, \beta < 1$ の例を考えるとよい。

# 連続時間 Markov 連鎖

この確率行列 $P$ で与えられる現象は、ある基本時間単位、例えば、1 年、1 日、1 時間、1 分、$\cdots$ など、現実的な時間間隔で状態が遷移する確率を与えている。
現実的な時間間隔が短くなった場合、例えば (技術革新などで) 今の 2 倍、3 倍、$\cdots$ とより多くの観測ができるようになった場合、さらには極限として連続時間で遷移が観測できるようになった場合、このマルコフ連鎖のモデルをどのように変えるのが良いだろうか？

# 連続時間 Markov 連鎖

現実的に同じ現象が見えるかは忘れて、数学的に取り扱いやすい方法は次のようになる。
小さな量 $\Delta t$ を問題を変化させた後の時間間隔とする。すなわち、もし 2 倍の観測が可能になった場合は $\Delta t = 1/2$ であるし、$n$ 倍の観測が可能になった場合は $\Delta t = 1/n$ となる。
(かなり大雑把に) 状態 $x_i \in S$ から状態 $x_j \in S$, $(x_i \neq x_j)$ への遷移する確率は元々与えられていた $P_{i,j}$ の $\Delta t$ 倍
$\tilde{P}_{ij} = \tilde{P}_{ij}(\Delta t) = \Delta t P_{ij}$ にする。
確率行列になるように対角成分で調整することにする、具体的には $P_{ii} = 1 - \sum_{j \neq i} P_{ij}$ で与えれれていたものを
$\tilde{P}_{ii} = \tilde{P}_{ii}(\Delta t) = 1 - \Delta t \sum_{j \neq i} P_{ij}$ で与える。
$\tilde{P}(\Delta t)$ で与えた確率行列は時間間隔 $\Delta t$ で見た時の遷移確率なので、元の確率と比較する場合は $1/\Delta t$ ステップ進めて
$\Delta t \times (1/\Delta t) = 1$ になった $(\tilde{P}(\Delta t))^{1/\Delta t}$ を比較するのが自然である。

2 状態の例で考えると、

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}, \qquad \tilde{P}(\Delta t) = \begin{pmatrix} 1-\Delta t\alpha & \Delta t\alpha \\ \Delta t\beta & 1-\Delta t\beta \end{pmatrix}$$

になる。

## 問題

$\tilde{P}(\Delta t)$ を対角化せよ。
さらに（$1/\Delta t$ が自然数になっていると思って）$(\tilde{P}(\Delta t))^{1/\Delta}$ を計算せよ。

# 連続時間 Markov 連鎖

$\Delta t$ は $1/n$ のように $1/$自然数　となるものを考えているように思えるが、数学としてみた場合、小さい量 $(\Delta t < 1)$ である限り意味を持つ。
表記の簡略化をする。
$\tilde{P}$ の定義の仕方から $P$ が与えられていると思うよりは、$P$ の非対角成分が与えらえていて、そこから ($P$ も含めて) $\tilde{P}$ を与えていると考えられるが、表記上は $P$ の対角成分も含めて与えられていて

$$Q = P - I, \qquad \tilde{P}(\Delta t) = \Delta t Q + I$$

と与えると統一的に表記できるのに加えて、$Q$ だけ与えれば全てが分かり使い勝手が良い。
一部の分野では「Q 行列」という言葉でこの行列を呼ぶこともある。(ただし連続時間マルコフ連鎖を取り扱う。)

# 連続時間 Markov 連鎖

## 問題

推移確率

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

に対応する $Q$ を与えよ

# 連続時間 Markov 連鎖

先にも書いたように $\tilde{P}(\Delta t)$ で与えた確率行列は時間間隔 $\Delta t$ で見た時の遷移確率なので、元の確率と比較する場合は $1/\Delta t$ ステップ進めて $\Delta t \times (1/\Delta t) = 1$ になった $(\tilde{P}(\Delta t))^{1/\Delta t}$ を比較するのが自然であり、これに $\tilde{P}(\Delta t) = \Delta t Q + I$ の表記を組み合わせると

$$(\tilde{P}(\Delta t))^{1/\Delta t} = (\Delta t Q + I)^{1/\Delta t}$$

と書けるが、 $Q, I$ が行列であることをあまり気にせず、$I \cong 1$ と思って $\Delta t \to 0$ の極限を考えれば

$$(\Delta t Q + I)^{1/\Delta t} \to \exp Q, \quad \Delta t \to 0$$
$$(\Delta t Q + I)^{\tau/\Delta t} \to \exp(\tau Q), \quad \Delta t \to 0$$

となるべきである。

# 連続時間 Markov 連鎖

これは一般的に、$A$ は対角化可能で

$$R^{-1}AR = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

と書けたとする。このとき $\exp tA$, は

$$\exp\{tA\} = R \begin{pmatrix} \exp(t\lambda_1) & & \\ & \ddots & \\ & & \exp(t\lambda_n) \end{pmatrix} R^{-1}$$

と定義することで正当化できる。

# 連続時間 Markov 連鎖

**問題**

推移確率

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}$$

に対応する $Q$ を用いて $\exp tQ$ を求めよ。

# 連続時間 Markov 連鎖

この $\exp(\tau Q)$ で与えられる行列は確率行列の性質を持つため、$\exp(\tau Q)_{ij}$ を時間 $\tau$ で $i$ から $j$ へ遷移する確率として連続時間マルコフ連鎖を考えることができる。
このように離散時間からの（形式的な）極限移行で連続時間マルコフ連鎖を与えたため、コンピュータによる（モンテカルロ）シミュレーションがよい近似になっていることは非常に分かりやすい。
一方でどのような運動をしているのかを考えるには、Poisson 分布を使った定義を用いた方が分かりやすいがここではこの 2 つの方法での定義は同じものを与えるということだけ述べるにとどめておく。

# 連続時間 Markov 連鎖

行列 $Q$ を与えるのと同値な方法を与える。
状態空間 $S$ は集合の要素 $\#S = r$ であり確率行列 $P$ を通して $Q = P - I$ が与えられていた。これに対して $S$ から実数への関数全体を $W$ と表すと自然に $r$ 次元の線形空間と考えられる。$W$ の元として次の $r$ 個の関数 $f_1, f_2, \ldots, f_r$ を考える
$x_i \in S$ $(i = 1, 2, \ldots, r)$ に対して

$$f_i(x_j) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

とする。このときこの $f_1, f_2, \ldots, f_r$ は線形空間 $W$ の基底となる。

# 連続時間 Markov 連鎖

線形空間 $W$ から $W$ への線形写像 $L$ を $Q$ を用いて

$$Lf(x_i) = \sum_{j \neq i} Q_{ij}(f(x_j) - f(x_i)), \qquad (i = 1, 2, \ldots, r)$$

と与える。この線形写像 $L$ の基底 $f_1, f_2, \ldots, f_r$ の表現行列を考えると $Q$ になる。
このため行列 $Q$ を与えずに、線形写像 $L$ を与えても良いことが分かる。
行列 $Q$ で書いたときに多くの行列成分が 0 になるような場合は非常に多く、そのような場合は、$L$ で書いた方が書きやすいため、$L$ を使うことが多くなる。
なお $L$ で書く場合には係数として文字 $Q$ を使わないで $c_{ij}$ や $c(i,j)$ のように $c$ を使うことが多い。

# 連続時間 Markov 連鎖

2 状態の例で考えると、

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}, \qquad Q = P - I = \begin{pmatrix} -\alpha & \alpha \\ \beta & -\beta \end{pmatrix}$$

に対応する $L$ を無理やり書くと

$$\begin{cases} Lf(1) = \alpha\{f(2) - f(1)\} \\ Lf(2) = \beta\{f(1) - f(2)\} \end{cases}$$

になるが、これくらい簡単な場合は $Q$ の方が分かりやすいかもしれない。

# 連続時間 Markov 連鎖

少し形式的ではあるが、単純ランダムウォークは

$$
\begin{aligned}
Lf(x) &= \frac{1}{2}\{f(x+1)+f(x-1)-2f(x)\} \\
&= \frac{1}{2}\{f(x+1)-f(x)\}+\frac{1}{2}\{f(x-1)-f(x)\}
\end{aligned}
$$

と単純に書けて便利である。多次元の場合も $i$ 軸方向の単位ベクトル $e_i$ を準備しておけば

$$
\begin{aligned}
Lf(x) &= \sum_{i=1}^{d} \frac{1}{2d}\{f(x+e_i)+f(x-e_i)-2f(x)\} \\
&= \sum_{i=1}^{d} \left[\frac{1}{2d}\{f(x+e_i)-f(x)\}+\frac{1}{2d}\{f(x-e_i)-f(x)\}\right]
\end{aligned}
$$

と簡単に書ける。

また第 1 回目に出てきたマルチンゲールが絡んだ問題を考えるにあたっては生成作用素は非常に使いやすい。