マルコフ連鎖モンテカルロ法

樋口さぶろお https://hig3.net

龍谷大学大学院 理工学研究科 数理情報学専攻

理論物理学特論　L07 (2022-11-02 Wed)

最終更新: Time-stamp: "2022-11-02 Wed 11:03 JST hig"

今日の目標

- マルコフ連鎖の極限分布の有無を判定できる
- マルコフ連鎖モンテカルロ法のメトロポリス・ヘイスティングス法を説明できる

略解 I

L06-Q1

Quiz 解答: マルコフ連鎖の転置推移確率行列

推移図略.

$$\begin{pmatrix} \frac{4}{7} & \frac{1}{7} & 0 & \frac{2}{7} \\ \frac{2}{7} & \frac{4}{7} & \frac{1}{7} & 0 \\ 0 & \frac{2}{7} & \frac{4}{7} & \frac{1}{7} \\ \frac{1}{7} & 0 & \frac{2}{7} & \frac{4}{7} \end{pmatrix}$$

$$\vec{p}(1) = M\vec{p}(0) = \begin{pmatrix} \frac{4}{7} \\ \frac{2}{7} \\ 0 \\ \frac{1}{7} \end{pmatrix}$$

L06-Q2

Quiz 解答: マルコフ連鎖の転置推移確率行列

略解 II

推移図略.

$$\begin{pmatrix} \frac{5}{7} & \frac{1}{7} & 0 & 0 \\ \frac{2}{7} & \frac{4}{7} & \frac{1}{7} & 0 \\ 0 & \frac{2}{7} & \frac{4}{7} & \frac{1}{7} \\ 0 & 0 & \frac{2}{7} & \frac{6}{7} \end{pmatrix}$$

$$\vec{p}(1) = M\vec{p}(0) = \begin{pmatrix} \frac{5}{7} \\ \frac{2}{7} \\ 0 \\ 0 \end{pmatrix}$$

ここまで来たよ

6 マルコフ連鎖

7 マルコフ連鎖モンテカルロ法
- 分布の時間発展
- 一般の場合: 可約なマルコフ連鎖, 非周期的でないマルコフ連鎖
- マルコフ連鎖モンテカルロ法

分布の時間発展 I

L07-Q1

Quiz(マルコフ連鎖の時間発展)

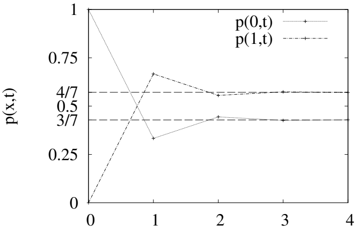
状態空間 $\{0,1\}$ 上のマルコフ連鎖を考える．転置推移確率行列は

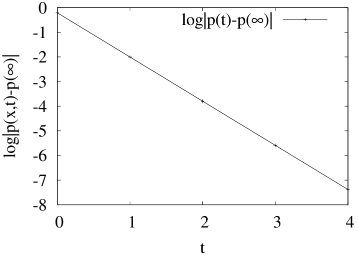
$$M = \begin{pmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{2} \end{pmatrix}.$$

1. 初期分布 $\vec{p}(0) = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ のとき $\vec{p}(1)$ を求めよう．
2. 定常分布をすべて求めよう．
3. 上の初期分布のとき $\vec{p}(t)$ を求めよう．
4. 上の初期分布のとき極限分布 $\vec{p}(\infty) = \lim_{t\to+\infty} \vec{p}(t)$ と収束の様子 $\log|\vec{p}(t) - \vec{p}(\infty)|$ を調べよう．
5. 初期分布 $\vec{p}(0) = \frac{1}{2}\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ のとき $\vec{p}(t)$ を求めよう．

定義 (極限分布)

(存在するなら) $\vec{p}(+\infty) = \lim_{t\to+\infty} \vec{p}(t)$ を極限分布という.




log|p(t)-p(∞)|
log|p(x,t)-p(∞)|
0
-1
-2
-3
-4
-5
-6
-7
-8
0
1
2
3
4
t

L07-Q2

Quiz(マルコフ連鎖の定常分布)

次の転置推移確率行列を持つ 状態空間 $\{0,1,2\}$ 上のマルコフ連鎖を考えよう.

$$M = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} & 0 \\ \frac{1}{4} & \frac{2}{4} & \frac{1}{4} \\ 0 & \frac{1}{4} & \frac{3}{4} \end{pmatrix}$$

なお, M の固有値固有ベクトルは $\lambda = 1, \frac{3}{4}, \frac{1}{4}, \quad \vec{u} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}$ であることを使ってよい.

1. 定常分布をすべて求めよう.
2. 初期分布 $\vec{p}(0) = \frac{1}{3}\begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$ のとき $\vec{p}(t)$ を求めよう.
3. 上のとき, $\log|\vec{p}(t) - \vec{p}(\infty)|$ の $t \to +\infty$ の振る舞いを求めよう.

単純な場合でのマルコフ連鎖の時間発展

状態空間 $S = \{0, 1, \ldots, m-1\}$ 上のマルコフ連鎖.
転置推移確率行列 M の固有値固有ベクトル $\lambda_i, \vec{u}_i$ $(i = 1, \ldots, m)$.

仮定

絶対値の大きさの順に並べると (第 1, 第 2,. . ., 第 m 固有値とよぶ) 次のようになっているとする.

$$1 = \lambda_1 > |\lambda_2| \geq \cdots \geq |\lambda_m| \geq 0.$$

このとき, $\vec{u}_1$ は確率ベクトルにとれて, 解は,

$$\vec{p}(t) = \vec{u}_1 + a_2\vec{u}_2\lambda_2^t + a_3\vec{u}_3\lambda_3^t + \cdots + a_m\vec{u}_m\lambda_m^t.$$

この単純な場合の性質

- 第 1 固有値は 1. 転置推移確率行列の固有値には, いつでも 1 が含まれることを示したのだった. 先週の証明
- 第 1 固有ベクトル $\vec{u}_1$ は 定常分布. ペロン-フロベニウスの定理
- 極限分布 $\vec{p}(+\infty) = \lim_{t\to+\infty} \vec{p}(t) = \vec{u}_1$. 自分の言葉でどうぞ

ちなみに, 極限分布が存在するなら, それは必ず定常分布. なぜなら 自分の言葉でどうぞ

定義 (定常分布)

$\vec{p} = M\vec{p}$ となる分布 $\vec{p}$ を M の定常分布 (stationary), 不変分布 (invariant) という. M は $\vec{p}$ を不変にするという.

確率変数列 $X(0), X(1), \ldots,$ が定常, という別の概念があるが, Markov 連鎖の場合, $X(0)$ が定常分布であることと同じになる.

この単純な場合の性質の続き

- 第 2 以降の固有値の絶対値が 1 より小なので 自分の言葉でどうぞ
  第 2 固有値の絶対値が小ささが, 極限分布への収束の速さを決める
- 第 2 以降の固有ベクトルは 自分の言葉でどうぞ

これは単純な場合で, 一般にはそうでない場合もある
次の節

ここまで来たよ

一般の場合 1: 可約なマルコフ連鎖 I

L07-Q3

Quiz(可約なマルコフ連鎖の定常状態)

次の転置推移確率行列を持つ 状態空間 $S = \{0, 1, 2\}$ 上のマルコフ連鎖を考える.

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

1. $\vec{p}(0) = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$ のとき時間発展 $\vec{p}(t)$ を求めよう.
2. $\vec{p}(0) = \frac{1}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ のとき時間発展 $\vec{p}(t)$ を求めよう.
3. 推移図を書こう.

Hint. 固有値 $\lambda = 1$(重解), $\frac{1}{3}$, 固有ベクトル $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$ 使用可.

定義 (既約 (irreducible) なマルコフ連鎖=可約でないマルコフ連鎖)

どの状態からどの状態へも, 確率 > 0 の矢印をたどって到達できるとき, マルコフ連鎖は (推移確率行列は) 既約であるという. 既約でないとき, 可約であるという.

可約なとき, 複数の定常分布が存在する. 極限分布は初期分布に依存する.
可約なマルコフ連鎖は, 既約なマルコフ連鎖に ‘分割して’ 考察すればよい.

一般の場合 2: 周期的な状態のあるマルコフ連鎖 I

L07-Q4

Quiz(マルコフ連鎖)

状態空間 $S = \{0, 1, 2\}$ 上のマルコフ連鎖を考える. 転置推移確率行列 M は次.

$$M = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

1. 定常分布をすべて求めよう.
2. 任意の初期分布は定常分布に近づくか考えよう.
3. 推移図を描こう.

Hint: $\lambda_j = \omega^j = (\frac{1+\sqrt{3}i}{2})^j \quad (j = 0, 1, 2)$ 固有ベクトル $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} s, \begin{pmatrix} 1 \\ \omega \\ \omega^2 \end{pmatrix} s, \begin{pmatrix} 1 \\ \omega^2 \\ \omega \end{pmatrix} s$

命題 (周期的な状態のあるマルコフ連鎖)

元の状態に戻る確率が, $k > 1$ 回目に初めてゼロでなくなるような状態 (周期的 (periodic) な状態) があるとき, マルコフ連鎖の転置推移確率行列には絶対値 1 の固有値が複数ある.

このとき, 初期分布によっては極限分布がないことがある.

マルコフ連鎖の分類

- 単純な場合=既約 (可約でない), かつ, 非周期的 (aperiodic, 周期的な状態がない)
- 一般には可約だったり周期的状態があったりする

既約で, 非周期的なとき, マルコフ連鎖はエルゴード的 (ergodic) であるという. 次に見るマルコフ連鎖モンテカルロ法では, マルコフ連鎖はエル

L07-Q5

Quiz(周期的なマルコフ連鎖の定常状態)

次の転置推移確率行列をもつ, 状態空間 $S = \{0, 1\}$ 上のマルコフ連鎖を考える.

$$M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

1. マルコフ連鎖の定常分布を求めよう.
2. $\vec{p}(0) = \frac{1}{3}\left(\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right)$ のとき $\vec{p}(t)$ を求めよう. 極限分布はある?
3. $\vec{p}(0) = \frac{1}{2}\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ のとき $\vec{p}(t)$ を求めよう. 極限分布はある?

ここまで来たよ

6 マルコフ連鎖

7 マルコフ連鎖モンテカルロ法
- 分布の時間発展
- 一般の場合: 可約なマルコフ連鎖, 非周期的でないマルコフ連鎖
- マルコフ連鎖モンテカルロ法

マルコフ連鎖モンテカルロ法の目的

目的 やりにくい確率関数, 確率密度関数 $f(x)$ にしたがう X の標本 (=乱数) を実際に抽出する．その標本期待値で, 母期待値を推定する．

> 確率の問題だけではない．例えば, 図形の面積は何かの確率変数の母期待値として書けることがある

前提

1. 計算機で, 一様分布 $\mathrm{U}(0,1)$ の独立同分布に「ほぼ」したがうような**擬似乱数** (pseudo-random number) は比較的容易に生成できる．
2. 任意の分布にしたがう確率分布にしたがう擬似乱数を作れるわけではない．
   - 作る工夫: 逆変換法, 棄却法, 個別の超絶技巧, $\cdots$．
   - 正規分布にしたがうものは一様分布から作れる．
3. $f(x)$ が定数倍だけ不明なことがある
   - $f(x) = \frac{1}{Z} g(x)$ で, 定数 $Z = \int_{-\infty}^{+\infty} g(x)\mathrm{d}x$ が計算できないことがある．
   - 同じことの別の言い方: $f(x)/f(y)$ は計算できるけど, $f(x)$ そのものは計算できない．

マルコフ連鎖モンテカルロ法 (MCMC)

どんな (定数 Z 不明の)$\frac{1}{Z}g(x)$ に対して, 定常分布が $\frac{1}{Z}g(x)$ になるような, うまいマルコフ連鎖 M を $f(x)$ から作って, シミュレートして, t における状態を取り出すことで擬似乱数を得る方法.

- うまい:$X(t)$ から $X(t+1)$ を確率的に定めるのは,$\mathrm{U}(0,1)$ 擬似乱数だけでできる

Listing 1: MCMC

```
double x=xinitial;
  for(n=0;n<サンプルサイズ;n++){
    for(t=0;t<待ち時間;t++){
      x=getnext(x); // U(0,1)一様乱数を使い,Mにしたがう
    }
  // 定常分布に近づいただろう
  printf("%f\n",x);
}
```

定義 (対称分布)

マルコフ連鎖 (の転置推移行列)M_{xy} に対して, 次の条件 (もちろん和はとってない) を満たす分布 p_x を, M の**対称分布** (symmtric) という.

$$\forall x, y \in S \quad M_{xy} p_y = M_{yx} p_x$$

- $P(X(0) = y$かつ$X(1) = x) = P(X(0) = x$かつ$X(1) = y)$ なので, 同時分布が $f_{\mathrm{X(0)X(1)}}(x, y) = f_{\mathrm{X(0)X(1)}}(y, x)$ だというのが '対称' の語源.
- 任意の 2 つの状態 x, y を結ぶ矢印は, 向きが逆のもの 2 本ある. ある時刻に, その上を通過する確率が同じ, という条件.

$\vec{p}$ が M の対称分布であることを, M は分布 p に関する**詳細つりあいの条件** (detailed balance condition) を満たすともいう.

命題

対称分布は不変分布 (定常分布) である.

p の次の時刻の分布 q を求める.

$$q_x = \sum_y M_{xy} p_y = \sum_y M_{yx} p_x = 1 \cdot p_x = p_x.$$

よって p_x は定常分布.

下心

p が定数倍を除いてわかっているとき, 詳細つりあいの条件を満たすマルコフ連鎖 M を作れれば, その定常分布 (運が良ければ極限分布) として p が現れるはず

メトロポリス・ヘイスティングス法

$p_x = p_x$ から詳細つりあいの条件を満たす M を作るうまい方法のひとつ.

メトロポリス・ヘイスティングス法 (Metropolis-Hastings)

$$M_{xy} = q_{xy} \min\{1, \frac{p_x q_{yx}}{p_y q_{xy}}\} + (\text{定数}) \times \delta_{xy}$$

定数は, 転置確率行列となるように調整できる.

q_{xy} 提案分布 自由に決めていい転置推移確率行列.
$q_{xy} = 1/|S|$ (全状態平等な提案) など, 対称行列が選ばれ $\min$ の中から消去されることが多い.
直観的には, q_{xy} で x が提案されたなら, p_x が p_y より大きいときは, かならず x に移る. 大きくないときは, 比に応じて確率的に移す.

メトロポリス・ヘイスティングス法

転置推移確率行列であること列の和が 1 になるように, (定数) で対角成分を調節できる.

対称であること ($x \neq y$ の場合に示す)

$p_x q_{yx} \geq p_y q_{xy}$ のとき

$$M_{xy} = q_{xy} \cdot 1, M_{yx} = q_{yx} \cdot \frac{p_y q_{xy}}{p_x q_{yx}}$$

よって, $M_{xy} p_y = q_{xy} p_y$. $M_{yx} p_x = q_{xy} p_y$.

$p_x < p_y$ のとき

$$M_{xy} = q_{xy} \cdot \frac{p_x q_{yx}}{p_y q_{xy}}, M_{yx} = q_{yx} \cdot 1$$

よって, $M_{xy} p_y = p_x q_{yx}. M_{yx} p_x = p_x q_{yx}$.

メトロポリス・ヘイスティングス法のアルゴリズム

Listing 2: MH in MCMC

```
//確率q(y,x)にしたがって提案yを選ぶ.
double getproposal(double y,double x);
double getuniform(void);//[0,1)一様擬似乱数

double getnext(double x){ //x:現在の状態
  int y; // 次の状態
  int yproposal;
  double u;
    yproposal=getproposal(y,x); // 提案.
    u=getuniform();
    if(u< min(1, (p(x)*q(y,x))/(p(y)*q(x,y)) ){
      y=yproposal; // 提案を採択
    } else {
      y=x; // 提案を棄却
    }
  return y; // 次の状態
}
```

混合分布でのパラメタ $\theta \in [0,1]$ のベイズ推定

理論物理学特論 (2022)L05

$$f_{\mathrm{Y}|}(y|\theta) = \theta f(y;\mu_0,\sigma_0) + (1-\theta) f(y;\mu_1,\sigma_1)$$

ベイズの定理から, θ の分布は,

$$f_{\Theta|\mathrm{Y}}(\theta|\{y_i\}) = \frac{\prod_{i=1}^{n}(\theta f(y_i;\mu_0,\sigma_0) + (1-\theta) f(y_i;\mu_1,\sigma_1))}{\int_{-\infty}^{+\infty}\prod_{i=1}^{n}(\theta f(y_i;\mu_0,\sigma_0) + (1-\theta) f(y_i;\mu_1,\sigma_1))\mathrm{d}\theta}.$$

記号がややこしい…
確率密度 $q_{\theta\theta'} = 1$,
分母を知らなくても, 採択確率に現れる比を

$$\frac{p_{\theta'}}{p_\theta}\frac{f(\theta')}{f(\theta)} = \frac{\prod_{i=1}^{n}(\theta' f(y_i;\mu_0,\sigma_0) + (1-\theta') f(y_i;\mu_1,\sigma_1))}{\prod_{i=1}^{n}(\theta f(y_i;\mu_0,\sigma_0) + (1-\theta) f(y_i;\mu_1,\sigma_1))}$$

で実行すれば, θ の事後分布からサンプルできる.