

MCMC によるベイズ推定

樋口さぶろお <https://hig3.net>

龍谷大学大学院 先端理工学部研究科 数理・情報科学コース

計算科学特論 L15(2025-01-15 Wed)

最終更新: Time-stamp: "2025-01-15 Wed 11:14 JST hig"

今日の目標

- マルコフ連鎖モンテカルロ法でベイズ推定できる
- PyMC パッケージを使える



ここまで来たよ

14 マルコフ連鎖モンテカルロ法と詳細つりあいの条件

15 MCMC によるベイズ推定

- 計算機でベイズ推定
- 二項モデルのベイズ推定への MCMC 適用例
- 周辺的话题

ベイズ推定はそう簡単にはできない…

Hoff ベイズ §4

計算科学 I(2024)L08

ベイズ推定の値は、観測値 $y_1, \dots, y_N$ から定まる、パラメタ θ の次の分布の母ナントカ.

$$f_{\Theta|Y}(\theta|y) = \frac{f_{Y|\Theta}(y|\theta)f_{\Theta}(\theta)}{\int_{-\infty}^{+\infty} f_{Y|\Theta}(y|\theta)f_{\Theta}(\theta)d\theta}$$

心配

- 知ってる分布になる？（共役事前分布ならなる）
- 知らない分布なら？
 - ▶ 分母の積分できる？
 - ▶ $E[\Theta|y] = \int_{-\infty}^{+\infty} \theta f_{\Theta|Y}(\theta|y)d\theta$ とかを自分で積分しないといけない.
- 事前分布 $f_{\Theta}(\theta)$ どう決める？

経験分布で代用

アイデア

母ナントカの計算が困難なら、 $f_{\Theta|Y}(\theta|y)$ からサイズ n の標本を抽出して、その分布 (**経験分布** (empirical distribution)) から推定すればいい。
モンテカルロ法 (MC=Monte-Carlo methods)

- 母期待値 $E[\Theta|y] \simeq$ 標本期待値 $\frac{1}{n}(\theta_1 + \dots + \theta_n)$
- 母比率 $\simeq$ 標本比率
- 母最頻値 $\simeq$ 経験分布のヒストグラムの最大階級値
- 母中央値 $\simeq$ 標本中央値

この n は、計算時間さえかければいくらでも大きくできる。(貴重な) 測定値 y_i の個数 N とは別。

- 知ってる分布になる場合、R や Python の `scipy.stats` で、`stats.dist.rvs(size=n)` で乱数を呼び出して取得できる。
- 知らない分布だったら？

心配タイプ1

- 知らない分布なら？

- ▶ サンプル抽出できる？ = 分布にしたがう乱数ある？ 書ける？
- ▶ 分母が計算できないなら？ 比率 $f_{|Y}(\theta|y)/f_{|Y}(\theta'|y)$ だけから乱数作れるといいな～

ひとつの解決策

マルコフ連鎖モンテカルロ法 (MCMC=Markov Chain Monte-Carlo methods)

マルコフ連鎖は確率過程 (stochastic process) の1分類.

⇒ 来週以降 今から

ここまで来たよ

14 マルコフ連鎖モンテカルロ法と詳細つりあいの条件

15 MCMC によるベイズ推定

- 計算機でベイズ推定
- 二項モデルのベイズ推定への MCMC 適用例
- 周辺的话题

二項モデル（複数のベルヌイ分布）

Hoff ベイズ §3.1

計算科学 I(2024)L03

値が不明なパラメタ θ の確率で表 $Y = 1$ が出るコインを N 回投げた結果を, $Y_1, \dots, Y_N = 0, 1$ とする.

Y_i は独立同分布にしたがう. $Y_i \stackrel{\text{i.i.d.}}{\sim} \text{B}(1, \theta)$ ($n = 1, p = \theta$ の二項分布). θ の確率分布を $y_1, \dots, y_N$ の情報で更新していきたい.

$N = 1$ のとき, $Y_1 = y_1$ を得る確率は (なぜか `given |` を使って),

$$f_{Y|}(y_1|\theta) = \theta^{y_1}(1 - \theta)^{1-y_1}.$$

一般の N のとき,

$$f_{Y|}(y_1, \dots, y_N|\theta) = \theta^{\sum_i y_i} (1 - \theta)^{N - \sum_i y_i}.$$

θ の分布

θ についての主観的確率（信念）は，連続型確率変数 Θ で表現できる．
何の先入観も持たないときの事前確率密度は $\Theta \sim U(0, 1)$ ，すなわち，一様分布

$$f(\theta) = \begin{cases} 1 & (0 \leq \theta \leq 1) \\ 0 & (\text{他}) \end{cases}$$

と考えられる．

このとき， $f_{Y|}(y_1, \dots, y_N | \theta)$ は条件付き確率であることが納得できる．

ベイズの定理より, $Y_1 = y_1, \dots, Y_N = y_N$ という結果を得たときの θ の事後確率密度は,

$$\begin{aligned} f_{|Y}(\theta|y_1, \dots, y_N) &= \frac{f(y_1, \dots, y_N, \theta)}{f_Y(y_1, \dots, y_N)} \\ &= \frac{f_{Y|} (y_1, \dots, y_N|\theta) f(\theta)}{f_Y(y_1, \dots, y_N)} \\ &= \frac{\theta^{\sum_i y_i} (1-\theta)^{N-\sum_i y_i} \cdot 1}{f_Y(y_1, \dots, y_N)} \quad (0 \leq \theta \leq 1) \end{aligned}$$

~~分子だけ見ると θ 依存性はわかるが, $Z =$ 分母まで正確に求めよう.~~

$$f_Y(y_1, \dots, y_N) = \int_0^1 p(y_1, \dots, y_N, \theta) d\theta = \int_0^1 \theta^{\sum_i y_i} (1-\theta)^{N-\sum_i y_i} \cdot 1 d\theta = \dots$$

~~ベータ分布, ベータ関数の迷宮へ~~ 計算科学 I(2024)L03

「厳密解」 $\Theta|y_1, \dots, y_N \sim \text{beta}(1 + \sum y_i, 1 + N - \sum y_i).$

MCMC の言葉では

マルコフ連鎖の状態 $x, y = \theta \in S = [0, 1]$.

ややこしい…上の Y は別の確率変数なので、ここでは状態を $x = \theta, y = \phi$ と書く.

提案分布の確率密度は $q_{\theta\phi} = 1$ などでもよい. 対称なのでメトロポリス法でいい.

$$M_{\theta\phi} = \min\left\{1, \frac{\theta^{\sum_i y_i} (1 - \theta)^{N - \sum_i y_i}}{\phi^{\sum_i y_i} (1 - \phi)^{N - \sum_i y_i}}\right\}.$$

上の M の計算には何も難しいところがない. ϕ は今のパラメタ, θ は提案された次のパラメタ. $y_1, \dots, y_N$ は観測結果のデータで, MCMC の実行中は単なる整数定数.

- $q_{\theta\phi} = 5\mathbf{I}_{[|\theta-\phi|=0.1]}(\theta)$ などとしてもいいが, 端 $y = 0, 1$ では対称でなくなるのでメトロポリス-ヘイスティングス法で.

サンプル Colab 参照

L15-Q1

Quiz(ポアソン分布のベイズ推定)

確率変数 Y は、未知のパラメタ θ のポアソン分布 $\text{Po}(\theta)$,

`Y=pymc.Poisson('Y',mu=q)` にしたがう.

Y のデータとして, 5, 7, 5, 9, 8 を得た.

θ の事前分布をパラメタ $a = 2, b = 4$ のガンマ分布 $\text{gamma}(2, 4)$,

`q=pymc.Gamma('q',alpha=2,beta=4)`

$$f(\theta) = \frac{1}{\Gamma(2)4^2} \theta e^{-\theta/4}$$

として, PyMC で, θ の事後分布のグラフを描き, 94% HDI を求めよう.

計算科学 I(2024)L15

Colab (Y のことを X_{obs} と書いてる),

計算科学 I(2024)L05

を参

考にしよう.

統計的プログラミング言語

- `scipy.stats`: Python のライブラリ
- R: 統計計算環境
- PyMC: Python に閉じて MCMC を実行するライブラリ
 - ▶ 抽象化されている．できることしかできない．
 - ▶ 宣言的
- PyStan, RStan: Python, R の中から Stan を実行するドライバ
- Stan: MCMC の実行に特化したソフトウェア
 - ▶ 詳細な指定．やり放題．
 - ▶ 手続き的？
- WinBUGS, OpenBUGS: MCMC の実行に特化したソフトウェア

ここまで来たよ

14 マルコフ連鎖モンテカルロ法と詳細つりあいの条件

15 MCMC によるベイズ推定

- 計算機でベイズ推定
- 二項モデルのベイズ推定への MCMC 適用例
- 周辺的话题

応用: 物理の人が考えた画像生成モデル

確率変数 $x = \{s_{ij}\}$. $w \times h$ の長形状に並んだ画素. 簡単のため $0 \leq s_{ij} \leq 2^8 - 1$ (グレイスケール).

画像が得られる確率 $f(x) = \frac{1}{Z} \prod_{(i,j),(k,l) \text{ が辺で連結}} g(s_{ij}, s_{kl})$.

辺の (pairwise な) 確率 $g(s, s')$ (s, s' が近い階調ほど大). (程度により画風が変わる).

MH の提案確率

$$q_{x'x} = \begin{cases} \frac{1}{wh} \frac{1}{2^8} & (x' \text{ は, } x \text{ の 1 画素 } s_{ij} \text{ だけを } 0 \leq s'_{ij} < 2^8 \text{ に変更したもの}) \\ 0 & (\text{他}) \end{cases}$$

MH の採択確率

$$q_{x'x} \times \min\{1, \bullet\} = q_{x'x} \times \min\left\{1, \frac{g(s'_{ij}, \text{隣})g(s'_{ij}, \text{隣})g(s'_{ij}, \text{隣})g(s'_{ij}, \text{隣})}{g(s_{ij}, \text{隣})g(s_{ij}, \text{隣})g(s_{ij}, \text{隣})g(s_{ij}, \text{隣})}\right\}$$

応用: シミュレーティッドアニーリング

確率的勾配降下法 SGD=Stochastic Gradient Descent の祖先

エネルギー ϕ , 外部パラメタ (逆温度) $\beta : 0 \rightarrow +\infty$.

Step1 $\theta_{[0]} :=$ 適当な初期値

Step2 改善

Step2.2 $\theta_{[t+1]} = \theta_{[t]} +$ ランダムな小さいベクトル

Step2.3 もし $\phi(\theta_{[t+1]}) \leq \phi(\theta_{[t]})$ なら Step2 終了, そうでなければ,

Step2.3.1 $R \sim U(0, 1)$ ($[0, 1)$ 連続型一様乱数)

Step2.3.1

$$R < \frac{e^{-\beta\phi(\theta_{[t+1]})}}{e^{-\beta\phi(\theta_{[t]})}}$$

なら $\theta_{[t+1]}$ を返す. そうでなければ元の $\theta_{[t+1]} := \theta_{[t]}$ を返す.

Step2.4 β を少し増やす.

Step3 実質的に '収束' したなら $\theta_{[t+1]}$ を出力して終了, そうでなければ, $t := t + 1$ して Step2 にもどる

