政治学方法論 II

第 13 回：マルコフ連鎖モンテカルロ法

矢内　勇生

法学部・法学研究科

2015 年 7 月 15 日

神戸大学

今日の内容

1. MCMC 入門
   - 事後分布のシミュレーション
2. MCMC
   - メトロポリス法 (Metropolis Algorithm)
   - 二項分布モデルへの応用
   - MH 法 (Metropolis-Hastings Algorithm)
3. マルコフ連鎖
   - マルコフ連鎖

ベイス推定の目的

事後分布を求める

- 事後分布

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} = \frac{p(y|\theta)p(\theta)}{\int p(y|\theta)p(\theta)d\theta}$$
$$\propto p(y|\theta)p(\theta)$$

- 問題：$p(y|\theta)p(\theta)$ は確率分布ではない！
- 正規化定数 $p(y) = \int p(y|\theta)p(\theta)d\theta$ を求める必要がある
- これまでの対処法
  - 自然共役事前分布を使って分布を特定する（解析）
  - グリッドを作って正規化する（数値計算 1）
  - 周辺分布と条件付き確率を使って母数を順番に抽出する（数値計算 2）
- より複雑な問題（事前分布が共役ではない、母数の範囲が不明、母数の数が多い等）に対応できない

非正規化密度 (unnormalized density)

- 求めたい分布（目標分布）：$p(\theta|y)$
- 目標分布のうち、正規化定数以外の部分 $q(\theta|y)$ は（容易に）計算可能であると仮定する
- $q(\theta|y)$：非正規化密度
- $\frac{q(\theta|y)}{p(\theta|y)} = c$：$c$ は y の関数（θ に関しては定数）
- 通常のベイズ分析における非正規化密度：

$$q(\theta|y) = p(y|\theta)p(\theta)$$

- 非正規化密度をうまく使って、目標分布（正規化密度、事後分布）を手に入れる

問題の設定

選挙区を「跳び」回る政治家 (Kruschke 2015: 146–156)

- 選挙キャンペーンために選挙区を巡回したい政治家
- 選挙区は複数の島 θ から成る（単純化のため、東西に並んでいると仮定）
- 政治家は、各期 $t = 1, 2, \dots$ にどの島を訪ねるか決める
- 政治家は $t+1$ 期には t 期にいた島の隣の島に移動する
- 政治家は、最終的には各島の人口に比例した回数だけ各島を訪れたい
- 政治家は、各島の人口を知らない
- 政治家は以下の 2 つの情報だけ得られる（過去に得た情報は忘れる）
  1. 現在いる島の人口
  2. 次に移動するつもりの島の人口

政治家の行動

1. 最初 ($t = 1$) に訪れる島を適当に決める
2. 公正なコインを投げて、表が出たら東の島、裏が出たら西の島に移動することを考える
3. 現在の島の人口 $P_{current} = q(\theta_c)$ と移動候補の島の人口 $P_{proposed} = q(\theta_p)$ の情報を得る
4. 移動するかどうか決める（確率的移動にはルーレットを使用）
   - $p(\theta_c) < p(\theta_p)$ のとき：移動する
   - $p(\theta_c) \geq p(\theta_p)$ のとき：確率 $p_{move} = \frac{p(\theta_p)}{p(\theta_c)}$ で移動し、$1 - p_{move}$ で現在の島にとどまる
5. 2〜4 を繰り返す

移動するかどうかを決める確率

- 次の期の状態（訪れる島）
  - 確率 p_M で移動

$$p_M = \min\left(\frac{p(\theta_p)}{p(\theta_c)}, 1\right) = \min\left(\frac{q(\theta_p)}{q(\theta_c)}, 1\right)$$

  - その他の場合は現在の島にとどまる
- p_M：受容確率
- 現在より密度が高い点が提案されれば必ず移動する
- 現在より密度が低い点が提案されても確率的に移動する

政治家の行動の結果

- 十分長い期間、前述の行動を繰り返す
- 人口に比例した回数だけ各島を訪れることになる！
- メトロポリス法 (Metropolis algorithm)
- $q(\theta)$ を正規化定数を除いた非正規化事後分布（尤度 $\times$ 事前分布）と考える：正規化定数を知らなくとも、事後分布をシミュレートできる
- ただし、以下が必要
    - ランダムに提案をする方法（ある確率分布に従う乱数を発生させることができる）
    - 提案され得るすべての母数の値について、$q(\theta_p)/q(\theta_c)$ が計算できる
    - p_M によって提案を受容するかどうか決めるために、一様分布 $\mathrm{U}(0,1)$ に従う乱数を発生させることができる

なぜメトロポリス法でうまくいくのか

- ある位置 θ と $\theta+1$ の間の移動を考える
- θ から $\theta+1$ に移動する確率：

$$p(\theta \to \theta+1) = 0.5\min\left(\frac{q(\theta+1)}{q(\theta)}, 1\right)$$

- $\theta+1$ から θ に移動する確率：

$$p(\theta+1 \to \theta) = 0.5\min\left(\frac{q(\theta)}{q(\theta+1)}, 1\right)$$

- 比を取ると

$$\frac{p(\theta \to \theta+1)}{p(\theta+1 \to \theta)} = \frac{q(\theta+1)}{q(\theta)}$$

- 2 つの位置の相対頻度に応じて行ったり来たりする
- すべての θ についてこれが成り立つ
- 結局、すべての θ が相対頻度に応じて選ばれる

問題の設定

- あるコインを 1 回投げて表が出る確率 θ を知りたい
- n 回コインを投げたところ、y 回表が出た
- サンプリングモデル：$y \sim \mathrm{Bin}(n, \theta)$

$$p(y|\theta) \propto \theta^y (1-\theta)^{n-y}$$

- 事前分布：$\theta \sim \mathrm{Beta}(\alpha, \beta)$

$$p(\theta) \propto \theta^{\alpha-1}(1-\theta)^{\beta-1}$$

- 事後分布：

$$p(\theta|y) \propto q(\theta|y) = \theta^{y+\alpha-1}(1-\theta)^{n-y+\beta-1}$$

メトロポリス法で事後分布をシミュレートする

- $\theta \in [0,1]$ は連続値
- 隣り合った数以外も提案できるようにする
- 提案分布としてランダムウォークを使う：

$$\theta_p = \theta_c + \varepsilon$$
$$\varepsilon \sim \mathrm{N}(0, \sigma^2)$$

- メトロポリス法
  1. 初期値 $\theta^{(1)} \in [0,1]$ を適当に選ぶ：ただし、$p(\theta^{(1)}|y) > 0$
  2. 乱数を使って移動先を提案する：$\theta_p \sim \mathrm{N}(\theta_c, \sigma^2)$
  3. 確率 $p_M = \min(q(\theta_p|y)/q(\theta_c|y), 1)$ で θ_p に移動、それ以外は θ_c に留まる
  4. 2, 3 を十分長い間繰り返す
- 実習：$n = 20, y = 14, \alpha = \beta = 1$ の場合について、メトロポリス法で事後分布を求めなさい

MH 法：メトロポリス法の一般化

t 時点での提案分布：$J_t(\theta^*|\theta^{(t-1)})$

1. 提案分布の対称性
   - メトロポリス法の提案：対称

$$J_t(\theta_a|\theta_b) = J_t(\theta_b|\theta_a)$$

   - MH 法：対称でなくてもよい
2. 受容確率
   - メトロポリス：$\min\left(p(\theta^*|y)/p(\theta^{(t-1)}|y), 1\right)$
   - MH:

$$\min\left(\frac{p(\theta^*|y)/J_t(\theta^*|\theta^{(t-1)})}{p(\theta^{(t-1)}|y)/J_t(\theta^{(t-1)}|\theta^*)}, 1\right)$$

独立 MH 法

- 通常の提案分布：条件付き分布

$$J_t(\theta^* | \theta^{(t-1)})$$

- 独立 MH 法の提案分布：互いに独立
- 例：二項分布の母数 θ を提案する
  - 通常の Metropolis (MH) 法：$\theta^* = \theta^{(t-1)} + \varepsilon$
  - 独立 MH 法：$\theta^* \sim \mathrm{N}(0.5, \sigma^2)$

マルコフ連鎖モンテカルロ法

- Markov chain Monte Carlo (MCMC)
- マルコフ連鎖を利用したモンテカルロシミュレーション

確率過程

- 確率過程：時間に従って確率的に状態が変化する系列

Ｔシャツの選択

- ３色のＴシャツを所有しており、毎日どの色を着るか決める
- 状態空間 $S = \{R, G, B\}$
- 時点（日）：$t = 1, 2, \ldots$
- t において選択する色：$X^{(t)}$
- Ｔシャツ選択の実現値の例

$x^{(1)}$	$x^{(2)}$	$x^{(3)}$	$x^{(4)}$	$x^{(5)}$	$x^{(6)}$	$\cdots$
R	R	G	B	B	G	$\cdots$

マルコフ連鎖

- $t+1$ 時点における選択の確率を、過去の選択の条件付き確率として表す

$$\Pr(X^{(t+1)} = j | X^{(1)} = s_1, X^{(2)} = s_2, \ldots, X^{(t)} = i)$$

- マルコフ連鎖：

$$\Pr(X^{(t+1)} = j | X^{(1)} = s_1, \ldots, X^{(t)} = i) = \Pr(X^{(t+1)} = j | X^{(t)} = i)$$

- マルコフ連鎖：$t+1$ 時点の状態が t 時点の状態のみに依存する時間が離散的な確率過程

状態の推移

- $K_{ij} = \Pr(X^{(t+1)} = j | X^{(t)} = i)$ とする
- $i = j$ なら K_{ij} は現在の状態に留まる確率
- 推移核 (transition kernel または 推移行列 [transition matrix])
  K：K_{ij} を i 行 j 列要素とする行列
- $\sum_j K_{ij} = 1$：必ずいずれかの状態に遷移する
- 推移行列の例（T シャツの色の選択）

$$K = \begin{bmatrix} 0.3 & 0.5 & 0.2 \\ 0.4 & 0.2 & 0.4 \\ 0.2 & 0.6 & 0.2 \end{bmatrix}$$

  - 赤いシャツの次の日：確率 0.3 で赤、0.5 で緑、0.2 で青
  - 緑のシャツの次の日：$\cdots$
  - 青いシャツの次の日：$\cdots$

状態の推移と推移行列

- 初期状態の確率分布を π_1 とする
  - 初日は絶対赤いシャツと決めている：$\pi_1 = (1,0,0)$
  - 初日は赤と青で五分五分：$\pi_1 = (0.5,0,0.5)$
- 第 2 期の状態の確率分布：$\pi_2 = \pi_1 K$
- 第 $t+1$ 期の状態の確率分布：$\pi_{t+1} = \pi_t K = \pi_1 K^t$
- $\pi = \pi K$ のとき π を推移核 K のマルコフ連鎖の定常分布 (stationary [invariance, equilibrium] distribution) と呼ぶ

マルコフ連鎖の既約性

- 有限な t 回の推移でマルコフ連鎖の状態 i から状態 j に到達する確率が正 ($K_{ij}^{t} > 0$) のとき：状態 i は j に到達可能である
- $K_{ij}^{t} > 0$ かつ $K_{ji}^{t'} > 0$ のとき：状態 i と j は互いに到達可能である
- マルコフ連鎖の状態空間 S のすべての要素が互いに到達可能なとき：マルコフ連鎖は既約的 (irreducible) である

マルコフ連鎖の周期性

- 状態 i から j に到達する推移回数に規則性があるか

じゃんけん

- 状態空間 $S = \{$ グ, チ, パ $\}$
- 以下の推移核 K をおく

$$K = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

- マルコフ連鎖： $\cdots$ グ, チ, パ, グ, チ, $\cdots$
- 3 回ごとに必ず同じ手：周期 3 のマルコフ過程

- 周期：連鎖の各状態が同じパタンで繰り返される推移回数
- 周期が 1 のとき：非周期的 (aperiodic) と呼ぶ

マルコフ連鎖の再帰性

- 再帰的（recursive）：状態空間の任意の集合 A は限りなく何度も訪問される
- $t \to \infty$ のとき、状態空間の任意の集合 A が訪問される回数の期待値は ∞ である
- ハリス再帰的 (Harris recurrent) ：状態空間の任意の集合 A が訪問される回数の期待値が ∞ になる確率が 1
- ハリス再帰的 $\Rightarrow$ どの初期値から始めても同じ定常分布に到達する

マルコフ連鎖

- 既約的で、非周期的で、再帰的なマルコフ連鎖：定常分布に収束する
- 定常分布の収束にはある程度時間がかかる：バーンイン期間が必要
- バーンイン期間の長さはわからない
- どう対処する？
- いつかは収束する（はず）
- 収束の判定基準 $\hat{R}$
  - 複数のマルコフ連鎖の分散を比較して収束を判定する
  - 収束している：どのマルコフ連鎖も定常分布を移動：連鎖間の分散はなくなる