ベイズ推定への改良マルコフ連鎖モンテカルロ法の適用

千田史哉　宗久知男　宗久保子
山梨大学　大学院医学工学総合教育部コンピュータメディア工学専攻

1．はじめに

本研究ではベイズ推定の過程に改良マルコフ連鎖モンテカルロ法を適用し未知母数の確率分布を求めることを目的とする．

ベイズ推定は観測された証拠をもとに原因となる事象を推定する手法である．これにマルコフ連鎖モンテカルロ法を使用することで，推定したい事象の確率分布をサンプリングにより求められるかどうかを検討する．

2．ベイズ推定

未知の確率値X（原因）と観測可能な事象D（結果）があるとき，未知の確率値Xの確率 $p(X)$ を事前分布，Dが観測された時のXの条件付確率 $p(X|D)$ を事後確率といい，この事後確率 $p(X|D)$ は以下の式が成立する．

$$p(X|D)=\frac{p(D,X)}{p(D)}=\frac{p(D|X)p(X)}{p(D)}$$

これをベイズの定理という．$p(D|X)$ は尤度と呼ばれ，尤度関数 $L(X|D)$ と置くとこの事後確率は

$$p(X|D)\propto L(X|D)p(X)$$

となり，尤度関数と事前確率に比例することが分かる．

3．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は確率分布からサンプリングを行う手法のことである．

マルコフ連鎖とは未来の状態 $x^{(t+1)}$ は現在の状態 $x^{(t)}$ にのみに依存して推移する動作を示す．この推移する確率を推移確率といい，これを行列表記したものを推移確率行列と呼ぶ．推移確率行列Pとすると，推移は初期状態 $\pi^{(1)}$ と推移確率行列Pで以下のように表される．

$$\pi^{(1)}P^{t}=\pi^{(t+1)}$$

ここで既約的，正再帰的かつ非周期的なマルコフ連鎖はエルゴード性を満たすと言われ，このときどのような初期状態から出発しても，定常分布に収束する．

また，マルコフ連鎖が定常分布に収束する十分条件として，詳細釣り合いの条件がある．この条件は状態空間に属する全てのxについて，

$$\pi(x^{(t)})p(x^{(t+1)}|x^{(t)})=\pi(x^{(t+1)})p(x^{(t)}|x^{(t+1)})$$

が満たされることをいう．

モンテカルロ法とは，乱数を用いた分布のサンプリングを行う手法であり，本研究ではマルコフ連鎖の状態を標本に，新しい標本をサンプリングし，その標本をサンプリングの前後で形成される分布の良し悪しを乱数に基づいて採択するメトロポリス・ヘイスティングスアルゴリズムを用いる．

また，本研究での改良マルコフ連鎖モンテカルロ法とは，連鎖時間に推移が依存するマルコフ連鎖を用いたモンテカルロ法のことである．推移確率が変更されることにより，採択されない確率を減らし，より正確に分布の推定を行うことが期待される．この改良マルコフ連鎖も上記条件を満たすことにより定常分布に収束することが証明されている．[1]

4．マルコフ連鎖モンテカルロ法を用いたベイズ推定

実際の実験では，1980年に行われたアメリカ大統領選挙の例を用いて改良マルコフ連鎖モンテカルロ法を適用したベイズ推定を行う．当時の選挙は主にロナルド・レーガン候補と現職であったジミー・カーター大統領の２名の候補者への投票結果を「人種」と「政治的思想」の違いにより区別した上で表１に示す．

表１．1980年アメリカ大統領選挙の得票数

	白人		非白人	
	レーガン氏	カーター氏	レーガン氏	カーター氏
とても自由主義	1	12	0	6
かなり自由主義	13	57	0	16
やや自由主義	44	71	2	23
どちらともいえない	155	146	1	31
やや保守主義	92	61	0	8
かなり保守主義	100	41	2	7
とても保守主義	18	8	0	4

説明変数は「人種」「政治的思想」であり，人種は白人ならば0，非白人ならば１と２値を符号化した．目的変数は「両者の得票率」とする．目

Application of modified Markov chain Monte Carlo method to Bayesian estimation

Fumiya Chida , Tomoo Munehisa , Yasuko Munehisa
: Yamanashi University Interdisciplinary Graduate School of Medicine and Engineering Computer Science and Media Engineering

的変数は２値変数であり，ここではレーガン氏の得票率を*p(x)*とし，カーター氏の得票率を*1-p(x)*とする．

表１は観測結果であるから，原因となる母数群βからなる分布の事後確率は観測結果Dからベイズの定理を用いて以下の式で表される．

$$p(\beta|D) \propto p(D|\beta)\,p(\beta)$$

母数群βはそれぞれ以下の母数に分けられる．

β_0　：切片

β_1　：人種の影響

β_{21} ～ β_{27}　：政治的思想の影響

（ β_{21}　：とても自由主義，　β_{22}　：かなり自由主義，・・・・，　β_{27}　：とても保守主義）

βの無情報事前分布をここでは，以下のように正規分布に従うよう設定する．

$$\beta \sim N(0,16^6)$$

またレーガン氏の得票率*p(x)*は説明変数の線形な合成変数　$Z=\beta_0+\beta_1 x_1+\beta_{21}x_{21}+\cdots+\beta_{27}x_{27}$
から，以下のように表される．

$$p(x)=\frac{\exp(Z)}{1+\exp(Z)}=\frac{1}{1+\exp(-Z)}$$

これらより，事後確率　$p(\beta|D)$　は

$$p(x)^{NR}(1-p(x))^{NC}\,p(\beta_0)\,p(\beta_1)\cdots p(\beta_{27})$$

に比例する．（NRはレーガン氏の得票数，NCはカーター氏の得票数）

ここでまず，マルコフ連鎖による山登り法で事後確率の最尤推定量を導くことを考える．

山登り法の手順は以下に示す通りである．

１）母数　$\beta_0,\beta_1,\cdots,\beta_{27}$　にランダムに初期値を与える．

２）母数　$\beta_i^{(t)}$　を一つ選び，マルコフ連鎖させ　$\beta_i^{(t+1)}$　に推移させる．

３）推移後の母数による事後確率　$p(\beta|D)^{(t+1)}$　が，推移前の事後確率　$p(\beta|D)^{(t)}$　より大きければ推移後の母数　$\beta_i^{(t+1)}$　を採択・更新し，小さければ推移前の母数　$\beta_i^{(t)}$　に戻す．

４）上記２）～３）をn回繰り返す．

本実験ではn=50,000回の繰り返しにより最尤推定量を求めた．実験結果は，数回この手順を試行した結果，最尤推定量　$\hat{\beta}$　は平均として以下のように求められた．

表２．最尤推定量の平均

β_0	0.08066	β_{24}	-0.03003
β_1	-2.88882	β_{25}	0.31146
β_{21}	-2.59664	β_{26}	0.84008
β_{22}	-1.57354	β_{27}	0.65457
β_{23}	-0.51804		

またこのとき，分布の最大値は全て同じ値であり，山登り法により事後分布の最大値が求まったことが分かる．

次に，マルコフ連鎖モンテカルロ法，メトロポリス・ヘイスティングアルゴリズムにより，事後分布を直接サンプリングする動作を以下に示す．

１）母数　$\beta_0,\beta_1,\cdots,\beta_{27}$　にランダムに初期値を与える．

２）母数　$\beta_i^{(t)}$　を一つ選び，マルコフ連鎖させ　$\beta_i^{(t+1)}$　に推移させる．

３）推移後の母数による事後確率　$p(\beta|D)^{(t+1)}$　が，推移前の事後確率　$p(\beta|D)^{(t)}$　より大きければ推移後の母数　$\beta_i^{(t+1)}$　を採択・更新し，小さければ下記５）の条件により分岐する．

５）事後確率の比が(0,1)の一様乱数ｒと比べ，

$$r < \frac{p(\beta|D)^{(t+1)}}{p(\beta|D)^{(t)}}$$

を満たすならば，採択・更新し，満たさない場合は推移前の母数　$\beta_i^{(t)}$　に戻す．

６）上記２）～５）をn回繰り返す．

本実験ではn=50,000回とし，うち10,000回をバーンイン（棄却）期間とした．

この方法でも山登り法と同様に，最尤推定量を導き出せたが，事後分布がサンプリング前後で改善されない場合も採択する可能性があり，山登り法のように最大値は安定して出力されなかったが，分布が推移前後で小さくなってもサンプリングする為，分布自体を推定していることが考えられる．

５．まとめ

マルコフ連鎖を用いた山登り法により，事後分布の最尤推定量・最大値を求めることができた．またマルコフ連鎖モンテカルロ法，メトロポリス・ヘイスティングアルゴリズムによりマルコフ連鎖を基にサンプリングを行い，分布を直接推定できることも検証できた．改良マルコフ連鎖モンテカルロ法によるベイズ推定を現在検討しており，これを適用することでより効率的に分布の推定を行えることが期待できる．

参考文献

[1]鈴木悠也：一様でない遷移確率を用いた焼き鈍し法，山梨大学卒業論文（2008）
[2]松原望：入門ベイズ統計
[3]豊田秀樹：マルコフ連鎖モンテカルロ法
[4]千田史哉：ベイズ定理を用いた情報研鑽における確率推論，山梨大学卒業論文（2009）