高次元でのマルコフ連鎖モンテカルロ法

大阪大学大学院基礎工学研究科，金融・保険教育研究センター，JST CREST　　鎌谷研吾

マルコフ連鎖モンテカルロ (MCMC) 法の高次元での有効性を解析する．$P_d(\mathrm{d}x) = p_d(x)\mathrm{d}x$ は $\mathbb{R}^d$ 上の確率分布とする．最も基本的な MCMC 法であるランダムウォーク型メトロポリス (RWM) 法は次の手続きで定義された．ある $x \in \mathbb{R}^d$ を出発点として以下を繰り返す：

Random walk Metropolis algorithm
- $x^* \leftarrow x + w,\ w \sim N_d(0, \sigma^2 I_d)$.
- $x \leftarrow x^*$ with probability $\alpha(x, x^*)$, ただし,

$$\alpha(x, x^*) = \min\left\{1, \frac{p_d(x^*)}{p_d(x)}\right\}.$$

上記のステップにより d 次元の確率変数列 $\{X_m^d\}_m$ が定義される．このマルコフ連鎖は $P_d(\mathrm{d}x)$ を不変分布として持つように出来ていて，例えば $p_d(x) > 0$ $(x \in \mathbb{R}^d)$ であればエルゴード的になり，下記右辺の積分が存在する限り大数の法則

$$\frac{1}{M}\sum_{m=1}^{M} f(X_m^d) \to \int f(x) P_d(\mathrm{d}x) \text{ a.s. } (M \to \infty) \tag{1}$$

が成り立つ．したがって右辺の積分の近似値として左辺の有限和を用いることが正当化されるわけである．

次元の高いベイズ推定では MCMC 法の利用が一般的である．しかし，高次元・大標本・モデルの複雑さはいずれも MCMC 法のパフォーマンスに悪影響を及ぼす．例として確率微分方程式

$$\mathrm{d}X_t = b(X_t, \beta)\mathrm{d}t + a(X_t, \alpha)\mathrm{d}w_t$$

の解となるプロセスの離散観測 $X_0, X_h, \ldots, X_{Nh}$ からパラメータ推定を考えよう．尤度が陽に書けないので擬似対数尤度

$$\begin{aligned} H_N(\alpha, \beta) = &- \frac{1}{2h}\sum_{n=1}^{N} A(X_{(n-1)h}, \alpha)^{-1}[(\Delta_n X - hb(X_{(n-1)h}, \beta))^{\otimes 2}] \\ &- \frac{1}{2}\sum_{n=1}^{N} \log\det A(X_{(n-1)h}, \alpha) \end{aligned}$$

の利用が一般的である．ここで $A = aa^T$ であり，$\Delta_n X = X_{nh} - X_{(n-1)h}$ である，しかし例えば以下の三次元の確率過程のモデルを見れば分かるように，H_N の計算はしばしば高コストで，従って一回の MCMC の繰り返しにかかる計算負荷が高くなる．

$$\begin{aligned} &\begin{pmatrix} \mathrm{d}X_t^1 \\ \mathrm{d}X_t^2 \\ \mathrm{d}X_t^3 \end{pmatrix} \\ &= \begin{pmatrix} -b_1 X_t^1 + 2\cos(X_t^2) + 2\sin(X_t^3) + b_2 \\ 4\sin(X_t^1) - b_3 X_t^2 + 3\cos(X_t^3) + b_4 \\ 4\cos(X_t^1) + 3\sin(X_t^2) - b_5 X_t^3 + b_6 \end{pmatrix}\mathrm{d}t \\ &\quad + \begin{pmatrix} a_1(1+\sin(a_3 X_t^2)) + \frac{a_4}{\sqrt{1+(X_t^1)^2 + a_2(X_t^3)^2}} & 0 & 0 \\ 0 & a_2(1+\sin(a_4 X_t^3)) + \frac{a_5}{\sqrt{1+(X_t^2)^2 + a_6(X_t^1)^2}} & 0 \\ 0 & 0 & a_3(1+\sin(a_5 X_t^1)) + \frac{a_6}{\sqrt{1+(X_t^3)^2 + a_4(X_t^2)^2}} \end{pmatrix}\mathrm{d}w_t \end{aligned}$$

計算負荷増大をいかに小さく抑えるかが重要である．ここでは高次元での計算負荷増大について解析する．RWM 法が困難なほどの高次元積分では，ラプラス変換を援用する方法や，並列化を用いる方法など様々な方法が存在する．本発表では少ない繰り返し回数で良い近似を与える次の MCMC 法を紹介する：

- $r \sim \mathrm{Gamma}(d/2, \|x-\mu\|^2/2)$,
- $x^* \leftarrow \mu + \rho^{1/2}(x-\mu) + (1-\rho)^{1/2} r^{-1/2} w, w \sim N_d(0, \sigma^2 I_d)$,
- $x \leftarrow x^*$ with probability $\alpha(x, x^*)$, ただし，

$$\alpha(x,y) = \min\left\{1, \frac{p(y)\|x-\mu\|^{-d}}{p(x)\|y-\mu\|^{-d}}\right\}.$$

ここで $\rho \in (0,1)$，$\mu \in \mathbb{R}^d$ は事前に選んでおくチューニングパラメータで，とくに μ の選択は結果に直結する．理論的には高次元漸近論やエルゴード性で正当化されるが，本発表では μ の選択を始めとする，具体的構成の工夫を述べる．本発表は論文 Kamatani [2014] にもとづいている．

本研究は JSPS 科研費 24740062 及び大阪大学未来研究ラボの助成を受けている．

参考文献

Kengo Kamatani. Efficient strategy for the Markov chain Monte Carlo in high-dimension with heavy-tailed target probability distribution. *Arxiv*, 2014. URL http://arxiv.org/abs/1412.6231.