高次元での高速な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，CREST, JST　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：

$$\begin{cases} x^* & \leftarrow & x + w,\ w \sim N_d(0, \sigma^2 I_d), \\ x & \leftarrow & x^* \text{ with probabilty } \alpha(x, x^*) \end{cases}$$

ただし，

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

このステップにより d 次元のマルコフ連鎖 $\{X_m^d\}_m$ 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．

次元が高くなるに連れ，(1) の右辺の数値積分は極めて困難になり，MCMC 法の意義が出る．しかし，高次元での計算負荷の増大は MCMC 法でも不可避である．これは $d \to \infty$ としたときの $\{X_m^d\}_m$ の振る舞いを見ることでその一端を説明できる．P_d の形状に強く依存するが，P_d の裾が軽ければ，MCMC の繰り返し回数は $O(d)$ 回必要であり，裾が重い場合は $O(d^2)$ 回必要になる．一回の $p_d(x)$ の値の計算負荷自体も d に依存するので，高次元で MCMC 法を適用する際は，計算負荷の増大を軽減するための工夫が不可欠である．

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．本発表では前述の繰り返し回数を改善するような次の MCMC 法を紹介する：

$$\begin{cases} r & \sim & \mathrm{Gamma}(d/2, \|x - \mu\|^2/2), \\ x^* & \leftarrow & \mu + \rho^{1/2}(x - \mu) + (1 - \rho)^{1/2} r^{-1/2} w, w \sim N_d(0, \sigma^2 I_d), \\ x & \leftarrow & x^* \text{ with probabilty } \alpha(x, x^*) \end{cases}$$

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

[Kamatani(2014)] 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`.