Spherical Hamiltonian Monte Carlo

竹田航太

2023 年 8 月 2 日

目次

概要

単位球面上の HMC を定式化することを目指す．さらに収束定理を示すのに必要な条件について整理する．セクション 2 から HMC について検討する．

1 単位球面上の分布とモンテカルロ法

1.1 von Mises-Fisher 分布

1.1.1 定義

Definition 1.1 (von Mises Fisher 分布 (vMF) [1])**.** $p \in \mathbb{N}$ とする．単位球面上のベクトル $\mu \in S^{p-1}$ と $\kappa > 0$ に対して，*von Mises Fisher* 分布は次の確率密度関数で与えられる．

$$f(x; \mu, \kappa) = C_p(\kappa)^{-1} \exp(\kappa \mu \cdot x)$$

ただし，$C_p(\kappa) = \frac{\kappa^{p/2-1}}{(2\pi)^{p/2} I_{p/2-1}(\kappa)}$ で I_ν は第 1 種修正ベッセル関数．

Remark 1.2. R^p 上の平均 $\mu \in S^{p-1}$ 分散 $\kappa^{-1} I$[*1]をもつ正規分布を S^p に制限したものが *von Mises Fisher* 分布である．μ が平均ベクトル，κ が集中度と呼ばれる．
vMF の確率密度関数は次のように導出される．$\mu \in S^p$ と $\kappa > 0$ に対して，$N(\mu, \kappa^{-1} I)$ の密度関数は

[*1] I は p 次元単位行列．

$x \in S^{p-1}$ のとき，

$$\exp\left(-\frac{\kappa}{2}(x-\mu)^2\right) = \exp\left(-\frac{\kappa}{2}(\|x\|^2 + \|\mu\|^2) + \kappa\mu\cdot x\right) = C\exp(\kappa\mu\cdot x)$$

と整理できる．ただし，最後の変形で $\|x\| = 1, \|\mu\| = 1$ を使い，定数 $C > 0$ を導入した．

Remark 1.3. $p = 3$ のときは

$$C_3(\kappa) = \frac{4\pi\sinh(\kappa)}{\kappa} = \frac{2\pi(e^{\kappa} - e^{-\kappa})}{\kappa}$$

となる．

Proof. 球座標表示に変換する．

$$\begin{aligned}
C_3(\kappa) &= \int_{S^2} f(x;\mu,\kappa)dx \\
&= \int_0^{2\pi}\int_0^{\pi} \exp(\kappa\cos(\theta))\sin(\theta)d\theta d\phi \\
&= \frac{2\pi}{\kappa}\int_{-\kappa}^{\kappa}\exp(t)dt \\
&= \frac{2\pi(e^{\kappa} - e^{-\kappa})}{\kappa}
\end{aligned}$$

□

1.1.2 サンプリング

$p = 3$ のとき単位球面上の von Mises Fisher 分布からのサンプリングを行う．$\mu \in S^2, \kappa > 0$ が与えられた時の $f(x;\mu,\kappa)$ からのサンプリングをする．

まず，平均ベクトルが $\mu_0 = (0,0,1)^\top$ の場合を考える．

Require: $\kappa > 0$

1: draw $v \sim \mathrm{Uni}(S^1)$
   draw $\phi \sim \mathrm{Uni}([0, 2\pi])$
   set $v = (\cos(\phi), \sin(\phi))$
2: draw $w \sim g(w) = c_\kappa^{-1}\exp(\kappa w)$ on $[-1, 1]$
   draw $y \sim \mathrm{Uni}[-1, 1]$
   set $w = Q_w(y) = \kappa^{-1}\log(e^{-\kappa} + (e^{\kappa} - e^{-\kappa})y)$
3: set $x = (\sqrt{1-w^2}v, w) \in \mathbb{R}^3$

ただし $c_\kappa = C_3(\kappa)/2\pi$ とおいた．また，$Q_w(y) = G_w^{-1}(y)$ であり，$G_w(t) = P(w \leq t) = c_\kappa^{-1}\kappa^{-1}(e^{\kappa t} - e^{-\kappa t})$ は $g(w)$ の分布関数．

次に，得られたサンプル $(x_n)_{n=1}^N \sim f(x;\mu = \mu_0, \kappa)$ を任意の $\mu \in S^2$ に対する $f(x;\mu,\kappa)$ に従うサンプルに変換する．$\mu_0 \to \mu$ と変換する球面上の回転行列 A_μ を用いて $(A_\mu x_n)_{n=1}^N$ とすれば，$f(x;\mu,\kappa)$ に従うサンプルが得られる．

A_μ の表示：平均ベクトル $\mu = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ が与えられた時，y 軸を中心に θ 回転させる行列 $R_y(\theta)$ と z 軸を中心に ϕ 回転させる行列 $R_z(\phi)$ を使って，$A_\mu = R_z(\phi)R_y(\theta)$ と書ける．（ただし，回転の方向は R_y は z 軸から x 軸に向かう方向，R_z は x 軸から y 軸に向かう方向とする．）

$$
\begin{aligned}
A_\mu &= R_z(\phi)R_y(\theta) \\
&= \begin{bmatrix} \cos\phi & -\sin\phi & 0 \\ \sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \\
&= \begin{bmatrix} \cos\theta\cos\phi & -\sin\phi & \sin\theta\cos\phi \\ \cos\theta\sin\phi & \cos\phi & \sin\theta\sin\phi \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}
\end{aligned}
$$

2　単位球面上の HMC

2.1　設定

$S^2 = \{\tilde{q} = (x, y, z) \in \mathbb{R}^3; \|\tilde{q}\| = 1\}$ に対して，S^2 上のポテンシャル $U : S^2 \to \mathbb{R}$ with $C_U = \int_{S^2} e^{-U(\tilde{q})} d\tilde{q} < \infty$ が与えられているとする．このとき，密度関数 $\pi_U(\tilde{q}) = C_U^{-1} e^{-U(\tilde{q})}$ に従ったサンプリングを S^2 上で行うことを考える．

計算で扱いやすくするため，球座標表示を導入する．$q = (\theta, \phi) \in [0, 2\pi) \times [0, \pi]$ に対して，

$$
\begin{cases} x &= \sin\theta\cos\phi \\ y &= \sin\theta\sin\phi \\ z &= \cos\theta \end{cases} \tag{2.1}
$$

と $\mathbb{R}^2$ から S^2 への写像を考える．この写像により R^3 の内積から S^2 に自然なリーマン計量が次のように与えられる．

$$
ds^2 (= dx^2 + dy^2 + dz^2) = d\theta^2 + \sin\theta^2 d\phi^2 = [d\theta, d\phi]\Sigma(q)^{-1} \begin{bmatrix} d\theta \\ d\phi \end{bmatrix} \tag{2.2}
$$

これは単位球面の第 1 基本形式に相当する．ただし，

$$
\Sigma(q) = \begin{bmatrix} 1 & 0 \\ 0 & \sin^{-2}\theta \end{bmatrix}
$$

であり，$\theta = 0, \pi$ のときに singular である．

2.2　運動量のサンプリングと運動エネルギー

2.2.1　Riemann metric

次に HMC を考える上で必要なのが各点 q における運動量のサンプリングのための条件付き分布（密度関数）$\pi_K(p|q)$ である．一般にはユーザーが自由に設定できるが，運動量の空間（接空間）T_qS^2 における metric を反映した正規分布を用いるのが自然である．今，T_qS^2 には (2.2) のように計量が入っており $\Sigma(q)$ を使って条件付き分布を

$$
\pi_K(p|q) = \frac{1}{2\pi|\Sigma(q)|} \exp\left(-\frac{1}{2}p^\top \Sigma(q)^{-1} p\right) \qquad (p \in T_qS^2 \sim \mathbb{R}^2)
$$

この分布から誘導される運動エネルギー $K(q,p)$ は

$$K(q,p) = \frac{1}{2} p^\top \Sigma(q)^{-1} p + \log(|\Sigma(q)|) + \mathrm{const} \tag{2.3}$$

である.

2.2.2 Euclid metric

Riemann metric を採用する必要はない.

3 Riemann mfd HMC

ref: Riemann mfd HMC, LD, BM; 意義と課題，実装 stormer Verlet

4 サンプリング例

vMF からのサンプリング

うまくいかなかったが球座標変換によるヤコビアンを考慮していなかったことがわかったので補正．すると，サンプリング中に log の引数がゼロになるエラーが発生した．2021/08/02：これは解決

2021/08/02：stan によるサンプリングでヤコビアンの補正を入れると積分値が理論値に収束した.

参考文献

[1] Kantil V. Mardia and Peter E. Jupp. *Directional Statistics*. WILEY SERIES ON PROBABILITY AND STATISTICS. JOHN WILEY and SONS, LTD, 2000.

[2] Gary Ulrich. Computer generation of distributions on the m-sphere. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 33(2):158–163, 1984.

[3] Sungkyu Jung. Generating von mises fisher distribution on the unit sphere (s 2). 2009.