



Statistical Computing

Chap. 4.1: Advanced MCMC

LIU, Ran

April 3, 2025

LIU, Ran - Department of Statistics @BNU

Summary

Introduction to Advanced MCMC

1.1 RJMCMC

1.2 LMC

1.3 HMC

1.4 NUTS

Reversible-jump MCMC (RJMCMC)

- RJMCMC 是 Metropolis-Hastings (MH) 算法的扩展，适用于模型维度可变的情形；
- 在贝叶斯变量选择、模型比较、混合模型、切换点检测等问题中广泛应用；
- 提出者：Peter Green (1995)，用于“跨维度”的 MCMC；
- 核心思想：通过引入 Jacobian 校正项，在 MH 接受率中保持“细致平衡” (detailed balance)。

Liu, Ran - Department of Statistics @BNU

MH 算法的接受率公式为：

$$\alpha(x, x') = \min \left\{ 1, \frac{\pi(x')q(x | x')}{\pi(x)q(x' | x)} \right\}$$

- 适用前提：当前状态 x 与候选状态 x' 必须在相同维度的空间中（即 $x, x' \in \mathbb{R}^d$ ）；
- 问题所在：若模型间参数个数不同（如变量选择问题中模型维度变化）， $x \in \mathbb{R}^d$, $x' \in \mathbb{R}^{d'}$ ，则：
 - 无法定义一个标准的提议密度 $q(x' | x)$ ，因为两者在不同空间上；
 - 无法直接比较 $\pi(x)$ 与 $\pi(x')$ ，因为定义在不同维度；
- 结论：MH 接受率不再满足“详细平衡”（detailed balance），导致采样错误。

解决方案：引入辅助变量与 Jacobian 修正项——这正是 RJMCMC 的核心思想。

目标是在目标分布 $\pi(x)$ 下构造满足详细平衡的马尔可夫链：

$$\pi(x)q(x' | x)\alpha(x, x') = \pi(x')q(x | x')\alpha(x', x)$$

引入辅助变量 $u \sim g(u)$ ，构造确定性变换：

$$(x', u') = h(x, u), \quad \text{可逆、可微}$$

变换前联合密度为：

$$\pi(x)g(u)$$

变换后联合密度（变量替换）为：

$$\pi(x')g'(u') \left| \frac{\partial(x', u')}{\partial(x, u)} \right|$$

Liu, Ran - Department of Statistics @BNU

RJMCMC 的接受率公式

RJMCMC 的接受概率带有 Jacobian 调整项：

$$\alpha(x, x') = \min \left\{ 1, \frac{\pi(x')g'(u')}{\pi(x)g(u)} \left| \frac{\partial(x', u')}{\partial(x, u)} \right| \right\}$$

- u, u' ：辅助变量（构造维度匹配）；
- $g(u)$ ：辅助变量的提议密度；
- $\left| \frac{\partial(x', u')}{\partial(x, u)} \right|$ ：Jacobian 行列式；
- 保证跨维度变换保持可逆与测度一致。

维度匹配原则

为了构造一个可逆变换 (diffeomorphism), 需要满足维度匹配条件:

$$\underbrace{d}_x + \underbrace{r}_u = \underbrace{d'}_{x'} + \underbrace{r'}_{u'}$$

- $x \in \mathbb{R}^d$, $x' \in \mathbb{R}^{d'}$;
- u, u' : 辅助变量, 用于“填补”维度差;
- 可以通过生成额外变量或丢弃变量来实现对称变换。

Liu, Ran - Department of Statistics @BNU

RJMCMC 算法步骤

- ① 当前状态: $x = (k, \theta_k)$, 模型为 k ;
- ② 提议切换到模型 k' , 并构造新参数 $\theta_{k'}$;
- ③ 使用辅助变量 $u \sim g(u)$, 构造变换: $(x', u') = h(x, u)$;
- ④ 计算接受率 $\alpha(x, x')$, 决定是否接受;
- ⑤ 若拒绝则保持原状态 x 。

LIU, Ran - Department of Statistics @BNU

Ref:

- Stochastic Computation Research Triangle @ University of Bristol
<https://people.maths.bris.ac.uk/~mapjg/slides/tdtut4.pdf>
- Reversible Jump MCMC (RJMCMC) @ University of Colorado Boulder
<https://www.colorado.edu/amath/sites/default/files/attached-files/rjmcmmc.pdf>

LIU, Ran - Department of Statistics @BNU

Summary

Introduction to Advanced MCMC

1.1 RJMCMC

1.2 LMC

1.3 HMC

1.4 NUTS

Langevin Monte Carlo (LMC)

朗之万方程 ($U(x) = -\log \pi(x)$):

$$dX(t) = -\nabla U(X(t))dt + \sqrt{2}\sigma dB_t \quad (1)$$

布朗运动的增量:

$$dB_t \sim \mathcal{N}(0, dt \cdot I) \quad (2)$$

考虑 $X(t)$ 的一个小增量:

$$x' - x = -\nabla U(x)dt + \mathcal{N}(0, 2\sigma^2 dt) \quad (3)$$

$$x' \sim \mathcal{N}(x - \nabla U(x)dt, 2\sigma^2 dt) \quad (4)$$

然后再算 x' 的接受率。

LMC 又称 Metropolis-adjusted Langevin algorithm (MALA), 有别于 Stochastic gradient Langevin dynamics (SGLD), 它具有接受拒绝来修正的步骤。

* 通过扩散极限理论 (diffusion limit theory) 分析, 得出在目标分布为 i.i.d. 正态的情况下, 最佳接受率为 0.574。

Gareth O. Roberts, Jeffrey S. Rosenthal, Optimal Scaling of Discrete Approximations to Langevin Diffusions, Journal of the Royal Statistical Society Series B: Statistical Methodology, Volume 60, Issue 1, January 1998, Pages 255–268, <https://doi.org/10.1111/1467-9868.00123>

Liu, Ran - Department of Statistics @BNU

Summary

Introduction to Advanced MCMC

1.1 RJMCMC

1.2 LMC

1.3 HMC

1.4 NUTS

Hamiltonian Monte Carlo

HMC(也称 Hybrid Monte Carlo) 通过在参数空间中模拟哈密顿动力学，生成一条马尔可夫链。它需要计算目标分布的梯度信息，因此对于具有可导概率密度函数的模型特别有效。HMC 通过使用梯度信息来指导采样过程，减少了无效的随机步骤，从而提高了采样效率。

引入一个辅助变量 M (momentum)，模拟粒子的运动来定义提案分布。

注意，HMC 无法处理含有离散变量的模型。(pyro 使用的时候，将离散变量积分掉了，使用的时候要小心。)

* 参考 [B 站]: 哈密顿力学初步

* 参考 [B 站]: The intuition behind the Hamiltonian Monte Carlo algorithm

* 参考: STAT 535: Statistical Machine Learning (UW) [Statistics @BNU](#)

* 在高维极限下，最佳接受率范围大约是 0.65–0.8，取决于 leapfrog 步数与步长。

Beskos, A., Pillai, N. S., Roberts, G. O., Sanz-Serna, J. M., & Stuart, A. M. (2013). Optimal tuning of the hybrid Monte Carlo algorithm. *Bernoulli*, 19(5A), 1501–1534.

LIU, Ran - Department of Statistics @BNU

Summary

Introduction to Advanced MCMC

1.1 RJMCMC

1.2 LMC

1.3 HMC

1.4 NUTS

No-U-Turn Sampler (NUTS)

No-U-Turn Sampler (NUTS) 是 Hamiltonian Monte Carlo (HMC) 算法的一种自适应扩展。它的核心思想是：在采样过程中自动决定轨迹长度，从而避免手动调整步数 L 带来的效率问题。

传统 HMC 需要预先设定模拟步数 L 。如果 L 太小，采样路径可能不足以探索参数空间；如果 L 太大，又可能造成采样效率下降，甚至反复“掉头”回到原地。

NUTS 的做法是：在模拟过程中检测是否出现了“掉头” (U-turn) 行为，并据此停止轨迹扩展。

NUTS 使用一种递归构建的二叉树 (binary tree) 结构来生成一条可能的采样路径。其过程如下：

- 从当前状态出发，随机初始化动量变量，开始构建轨迹。
- 每个节点表示一个状态（位置和动量），使用 leapfrog 积分器向前或向后扩展。
- 在每一轮中，树的深度增加一层，轨迹在两个方向上指数增长（例如第一轮构造 1 步，第二轮构造 2 步，第三轮构造 4 步，以此类推）。
- 在每一步扩展后，检查是否出现 “U-turn”：即新扩展的节点是否“掉头” 朝向已有轨迹的起点。

LiU, Ran - Department of Statistics @BNU

掉头的判定方法通常是：令 q^- 和 q^+ 分别表示当前轨迹的两端， p 是当前动量，则判断：

$$(q^+ - q^-) \cdot p < 0$$

如果该内积小于零，说明采样路径已经开始向原方向“掉头”，此时应终止轨迹扩展。这一判断可以在二叉树的构建过程中递归进行。

最终，NUTS 从构建好的轨迹中以一定概率选取一个状态作为采样结果，从而保持细致平衡。

* 参考 [B 站]：NUTS 算法介绍 by Michael Betancourt

LIU, Ran - Department of Statistics @BNU

可视化 MCMC 各种采样算法：

- MCMC Implementation (Github)

LIU, Ran - Department of Statistics @BNU