采样——MCMC

Mia Feng

2018 年 4 月 21 日

1 概述

MCMC：粗暴的采样模拟方式，用于模拟直接计算困难的分布。用于采样，数值积分等等。

求解目标：用多次采样得到的频率分布近似原概率分布。即本来对复杂的 $f(x)$ 做积分，但是因为 $f(x)$ 比较复杂所以显式积分困难。迂回方法是构造统计量 $\frac{f(x)}{p(x)}$，通过对 $x \sim p(x)$ 进行采样，求取统计量 $\frac{f(x)}{p(x)}$ 的期望得到数值积分值。

$$\theta = \int_a^b f(x)\, dx = \int_a^b \frac{f(x)}{p(x)} p(x) \approx \frac{1}{n} \sum_{i=0}^{n-1} \frac{f(x_i)}{p(x_i)} \tag{1}$$

求解思路：微积分思想（recall：学习微积分的时候，用无数个划分的小矩形的面积来近似面积，但是当时的小矩形是来自均匀分布的）。实际上，来自均匀分布的可能性很小，此时需要考虑对复杂分布如何模拟采样，一旦完成对复杂分布的描述就可以完成数值积分。

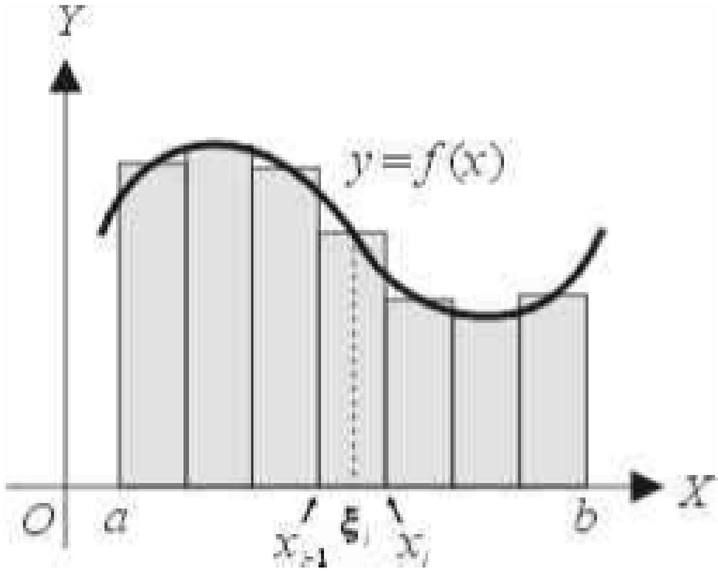

图 1: 矩形近似

求解方法：Markov Chain，蒙特卡洛积分，Metroplis-Hasting，Gibbs。
思维导图：详见 MCMC.xmind

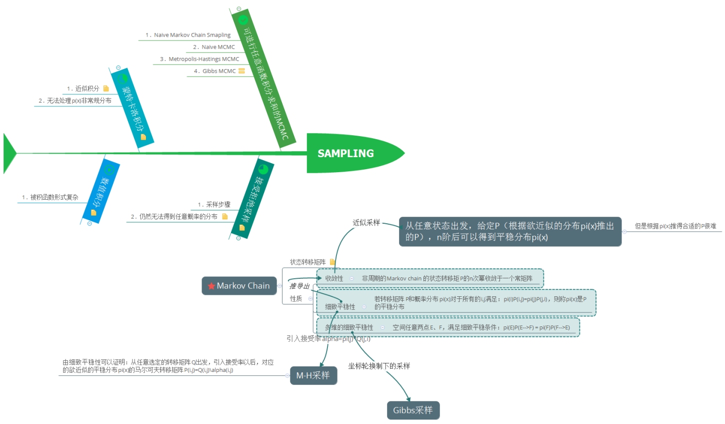

图 2: 思维导图

1.1 基本概念

马尔可夫矩阵的收敛性 可以参考 MIT 的线性代数课程中对马尔可夫矩阵的讲述。马尔可夫矩阵中各元素大于 0 且小于 1，而且矩阵是对称矩阵。马尔可夫矩阵的特征值中有一个为 1，其余都是比 1 小的正数，所以马尔可夫矩阵的 n 次幂收敛至一个常数，这也是为什么马尔科夫链一定会收敛，最终可以模拟一个平稳分布的原因。

马尔科夫链的细致平稳条件 如果非周期马尔科夫链的状态转移矩阵 P 和概率分布 $\pi(x)$ 对于所有的 i, j 满足

$$\pi(i) P(i, j) = \pi(j) P(j, i) \tag{2}$$

则称概率分布 $\pi(x)$ 是状态转移矩阵 P 的平稳分布。所以，利用对应的状态转移矩阵 P，就可以模拟平稳复杂分布 π。但是对于任意的平稳分布 π，P 的构造比较困难。MCMC 采用迂回的方式解决了这个问题。

多维数据的马尔可夫链的细致平稳条件 平面上任意两点 E, F，满足细致平稳条件

$$\pi(E) P(E \to F) = \pi(F) P(F \to E) \tag{3}$$

取上一状态的条件概率分布即可作为马尔科夫链的状态转移概率。

接受——拒绝采样 当公式 (1) 中的 $p(x)$ 不是常见分布时，无法根据 $p(x)$ 对 x 直接进行采样。此时接受——拒绝采样可以完成对 x 的模拟采样。选定提议分布 $q(x)$，如高斯分布等。先根据提议分布 $q(x)$ 采样得到一个样本 x_0，其对应的实际的概率是 u_0。再从均匀分布 $(0, kq(x_0))$ 中采样得到一值 u，若 $u < u_0$，接受这次采样值 x_0，反之拒绝。其中,$k, q(x)$ 的选取要确保 $kq(x) \geq p(x)$。当然，要满足这个条件比较困难，MCMC 中 Metropolis-Hastings 方法解决了这个问题。

马尔科夫链采样

MCMC 采样——MH 采样

MCMC 采样——Gibbs 采样

2 算法实现

注意实现时取了拉普拉斯平滑，见公式 (4)，且为了防止下溢取对概率值取了对数。[**?** **?**]

3 Implementation