

Markov Chain Monte Carlo 03 Metropolis Hastings Sampling

Chen Gong

01 January 2020

上一节中我们讲解了 Detailed Balance, 这是平稳分布的充分必要条件。Detailed Balance 为:

$$\pi(x)P(x \mapsto x^*) = \pi(x^*)P(x^* \mapsto x) \quad (1)$$

这里的 $P(x \mapsto x^*)$ 实际上就是条件概率 $P(z^*|x^*)$, 这样写只是便于理解。

首先, 我们需要明确一点, 我们要求的是后验概率分布 $P(Z)$, 也就是我们推断问题的核心目标。我们求 $P(Z)$ 主要是为了求在 $P(Z)$ 概率分布下的一个相关函数的期望, 也就是:

$$\mathbb{E}_{P(Z)}[f(Z)] \approx \frac{1}{N} \sum_{i=1}^N f(z^{(i)}) \quad (2)$$

而我们是通过采样来得到 $P(Z) \sim \{z^{(1)}, z^{(2)}, \dots, z^{(N)}\}$ 样本点。 $\pi(x)$ 是最终的平稳分布, 可以看成我们这里的 $P(Z)$, 下面的问题就是求出概率转移矩阵 P_{ij} , 才能满足 Detailed Balance 条件。知道了上面的条件以后, 我们每次这样进行采样, $x_1 \sim P(x|x_1)$, $x_2 \sim P(x|x_1)$, $x_3 \sim P(x|x_2)$, \dots , x_N 。最终就可以得到我们想要的 N 个样本。

1 Proposal Matrix

那我们怎么来找这个状态转移矩阵 P_{ij} 呢? 首先我们可以随机一个状态转移矩阵 Q_{ij} , 也就是 Proposal Matrix。

那么肯定是:

$$P(Z)Q(Z^*|Z) \neq P(Z^*)Q(Z|Z^*) \quad (3)$$

那么我们就想办法找到 Q_{ij} 使得:

$$P(Z)Q(Z^*|Z) = P(Z^*)Q(Z|Z^*) \quad (4)$$

那么, 我们怎么来解决这个问题呢? 我们可以在左右两边乘上一个因子来解决这个问题。也就是,

$$P(Z) \underbrace{Q(Z^*|Z)\alpha(Z^*, Z)}_{P(Z \mapsto Z^*)} = P(Z^*) \underbrace{Q(Z|Z^*)\alpha(Z, Z^*)}_{P(Z^* \mapsto Z)} \quad (5)$$

而 $\alpha(Z, Z^*)$ 定义为接收率, 大小为:

$$\alpha(Z, Z^*) = \min \left(1, \frac{P(Z^*)Q(Z|Z^*)}{P(Z)Q(Z^*|Z)} \right) \quad (6)$$

这样定义就行了？就可以满足 Detailed Balance 吗？我们可以证明一下，

$$\begin{aligned}
P(Z)Q(Z^*|Z)\alpha(Z, Z^*) &= P(Z)Q(Z^*|Z) \min\left(1, \frac{P(Z^*)Q(Z|Z^*)}{P(Z)Q(Z^*|Z)}\right) \\
&= \min(P(Z)Q(Z^*|Z), P(Z^*)Q(Z|Z^*)) \\
&= P(Z^*)Q(Z|Z^*) \min\left(\frac{P(Z)Q(Z^*|Z)}{P(Z^*)Q(Z|Z^*)}, 1\right) \\
&= P(Z^*)Q(Z|Z^*)\alpha(Z^*, Z)
\end{aligned} \tag{7}$$

那么我们就成功的证明了：

$$\underbrace{P(Z)Q(Z^*|Z)\alpha(Z, Z^*)}_{P(Z \mapsto Z^*)} = \underbrace{P(Z^*)Q(Z|Z^*)\alpha(Z^*, Z)}_{P(Z^* \mapsto Z)} \tag{8}$$

所以， $P(Z)$ 在转移矩阵 $Q(Z|Z^*)\alpha(Z^*, Z)$ 下是一个平稳分布，也就是一个马尔可夫链，通过在这个马尔可夫链中采样就可以得到我们的相应的数据样本点了。实际上这就是大名鼎鼎的 Metropolis-Hastings 采样法。

2 Metropolis-Hastings Sampling

第一步，我们从一个均匀分布中进行采样， $u \sim U(0, 1)$ ；

第二步，从 $Q(Z|Z^{(i-1)})$ 中进行采样得到样本点 Z^* ；

第三步，计算接受率， $\alpha = \min\left(1, \frac{P(Z^*)Q(Z|Z^*)}{P(Z)Q(Z^*|Z)}\right)$ 。注意，这里的 $P(Z) = \frac{\hat{P}(Z)}{Z_p}$ 。其中 Z_p 指的是归一化因子，几乎非常难以计算，所以一般是未知的。而 $\hat{P}(Z) = \text{likelihood} \times \text{prior}$ 。所以这里的 $P(Z)$ 和 $P(Z^*)$ 就是 $\hat{P}(Z)$ 和 $\hat{P}(Z^*)$ 。由于归一化因子被抵消了，干脆就直接写成了 $P(Z)$ 和 $P(Z^*)$ 。

第四步，如果 $u \leq \alpha$ 时 $Z^i = Z^*$ ，不然 $Z^i = Z^{(i-1)}$ 。

这样执行了 N 次，就可以得到 $\{Z^{(1)}, Z^{(2)}, \dots, Z^{(N)}\}$ 个样本点。