

Markov Chain Monte Carlo 04 Gibbs Sampling

Chen Gong

02 January 2020

如果我们要向一个高维的分布 $P(Z) = P(Z_1, Z_2, \dots, Z_N)$ 中进行采样。那么我们怎么来进行采样呢？我们的思想就是一维一维的来，在对每一维进行采样的时候固定住其他的维度，这就是 Gibbs Sampling。

我们首先规定一个 z_{-i} 是去除 z_i 后的序列， $\{z_1, z_2, \dots, z_{i-1}, z_{i+1}, \dots, z_N\}$ 。

1 A Example

假设 t 时刻，我们获得的样本为 $z_1^{(t)}, z_2^{(t)}, z_3^{(t)}$ 。

那么 $t+1$ 时刻，我们的采样顺序为：

$$\begin{aligned} z_1^{(t+1)} &\sim P(z_1 | z_2^{(t)}, z_3^{(t)}) \\ z_2^{(t+1)} &\sim P(z_2 | z_1^{(t+1)}, z_3^{(t)}) \\ z_3^{(t+1)} &\sim P(z_3 | z_1^{(t+1)}, z_2^{(t+1)}) \end{aligned} \quad (1)$$

从这个例子中，我们应该可以大致理解固定其他的维度然后进行一维一维采样的意思了。而实际上 Gibbs 是一种特殊的 MH 采样，为什么呢？我们来证明一下。

2 接受率 α 的计算

我们首先回顾一下，MH 采样的方法。我们的目的是从 $Q(Z|Z^{(t)})$ 中采样获得 Z^* ，然后计算接受率

$$\alpha = \min \left(1, \frac{P(Z^*)Q(Z|Z^*)}{P(Z)Q(Z^*|Z)} \right) \quad (2)$$

首先我们来看 $Q(Z|Z^{(t)})$ ：

$$Q(Z|Z^{(t)}) = Q(Z_i, Z_{-i} | Z_i^{(t)}, Z_{-i}^{(t)}) \quad (3)$$

假设我们现在是在对第 i 维进行采样，我们只要关注 $P(Z_i^* | Z_{-i})$ 。所以，我们可以得到： $Q(Z|Z^{(t)}) = P(Z_i^* | Z_{-i}^{(t)})$ 。

已经成功的将 $Q(Z|Z^{(t)})$ 做了等价转换以后。那么我们想要的 α 可以被我们成功的转换成如下的形式：

$$\alpha = \min \left(1, \frac{P(Z_i^* | Z_{-i}^*) P(Z_{-i}^*) P(Z_i | Z_{-i}^*)}{P(Z_i | Z_{-i}) P(Z_{-i}) P(Z_i^* | Z_{-i})} \right) \quad (4)$$

计算到了这里，我们还是不好进行计算，上面和下面好像还是不好消除。如果我们可以得到 Z_{-i}^* 和 Z_{-i} 之间的关系就好了。下面我们会得出一个重要的结论来帮助我们计算 α 的具体值。首先我们来举一个例子：

那么假设当 $t = 1$ 的时刻，有一个样本为： $Z_1^{(1)}, Z_2^{(1)}, Z_3^{(1)}$ 。

当 $t = 2$ 的时刻，我们假设先对第一维进行采样就可以得到： $Z_1^{(2)}, Z_2^{(1)}, Z_3^{(1)}$ 。

很显然 $Z_2^{(1)}, Z_3^{(1)}$ 根本没有发生变化。我们可以得到 $Z_{-1} = Z_{-1}^*$ 。也就是在 Gibbs 采样时，采样前后只关注于一个维度，其他的维度我们都没有关注到。所以就可以得到结论：

$$Z_{-i} = Z_{-i}^* \quad (5)$$

那么，我们把这个结论代入到公式 (4) 中，就可以得到：

$$\alpha = \min \left(1, \frac{P(Z_i^*|Z_{-i}^*)P(Z_{-i}^*)P(Z_i|Z_{-i}^*)}{P(Z_i|Z_{-i}^*)P(Z_{-i}^*)P(Z_i^*|Z_{-i}^*)} \right) = 1 \quad (6)$$

那么计算出接受率为 1，也就是每次都必定被接受。所以，每次从 $Q(Z|Z^{(t)}) = P(Z_i^*|Z_{-i})$ 中进行采样得到 Z_i^* 即可，一维一维的进行采样就可以采到整个高维的分布，各个维度上的样本。

所以，解释到了这里，大家基本就可以知道 Gibbs Samplings 是 $\alpha = 1$ 的 MH Sampling 的意义了。在 Gibbs Sampling 中 $\alpha = 1$ ，而且状态转移矩阵 $Q(Z|Z^{(t)}) = P(Z_i^*|Z_{-i}^{(t)})$ ，所以 Gibbs Sampling 就是把目标分布 P 对应的条件概率当作状态转移分布 Q 。

这里我们需要额外提醒一下，使用 Gibbs Sampling 是有使用前提的，也就是固定其他维度后的一维分布时方便进行采样的，如果固定其他维度的时候得到的一维分布仍然是非常难进行采样的，那么使用 Gibbs Sampling 也是没有用的。