**读书报告**

61518424 王贵涛

1. **问题与解答**

我提出问题：

1. 为什么独立抽样实现简单，但收敛速度慢？

讨论结果：我觉得如果是独立抽样的话，就是完全就是q(x, x')=q(x')，即建议分布只考虑了状态x'，并没有考虑x，所以这样就和接受拒绝法没有什么区别了，所以拒绝率可能会比较高，导致可能有很多时候位置是不动的，因此收敛速度可能就会比较慢。相较之下第一种方法当两者相近时a值较大因此接受概率较高，就会更容易转移，从而收敛。

别人提出的问题：

1. 吉布斯抽样比起直接使用Metropolis-Hastings算法的优势是什么？

我的解答：吉布斯抽样可以认为是Metropolis-Hastings算法的特殊情况，但是更容易实现，因而被广泛使用。它计算起来相对于MH算法更为简便。

1. P379贝叶斯公式的满概率分布正比于三个概率的式子应该如何推导出来？

我的解答：

1. **下周计划安排**

准备期末复习。

1. **读书收获**

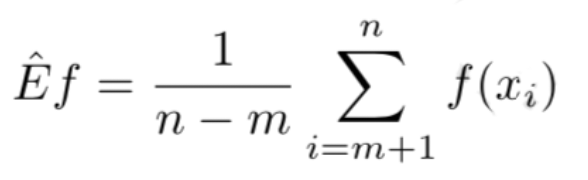
马尔可夫链蒙特卡罗法更适合于随机变量是多元的、密度函数是非标准形式的、随机变量各分量不独立等情况。

假设多元随机变量x，满足，其概率密度函数为p(x)，f(x)为定义在上的函数。目标是获得概率分布p(x)的样本集合，以及求函数f(x)的数学期望。

**基本思想**

在随机变量x的状态空间S上定义一个满足遍历定理的马尔可夫链，使其平稳分布就是抽样的目标分布p(x)。然后在这个马尔可夫链上进行随机游走，每个时刻得到一个样本。根据遍历定理，当时间趋于无穷时，样本的分布趋近平稳分布，样本的函数均值趋近函数的数学期望。

所以，当时间足够长时（时刻大于某个正整数m)，在之后的时间（时刻小于等于某个正整数n，n＞m）里随机游走得到的样本集合就是目标概率分布的抽样结果。得到的函数均值（遍历均值）就是要计算的数学期望值：



到时刻m为止的时间段称为燃烧期。

常用的马尔可夫链蒙特卡罗法有Metropolis-Hastings算法、吉布斯抽样。

**Metropolis-Hastings算法**

假设要抽样的概率分布为p(x)。Metropolis-Hastings算法采用转移核为p(x, x')的马尔可夫链：



其中q(x, x’)和α(x, x’)分别称为建议分布（proposal distribution）和接受分布(acceptance distribution)

转移核为p(x, x’)的马尔可夫链上的随机游走以以下方式进行：

如果在时刻(t-1)处于状态x，即xt-1=x，则先按建议分布q(x, x’)抽样产生一个候选状态x’，然后按照接受分布α(x, x’)抽样决定是否接受状态x’。

以概率α(x, x’)接受了x’，决定时刻t转移到状态x’，而以概率1-α(x, x’)拒绝x’，决定时刻t仍停留在状态x。

**吉布斯抽样**

吉布斯抽样是马尔可夫链蒙特卡罗法的常用算法吉布斯抽样，用于多元变量联合分布的抽样和估计。

其基本做法是，从联合概率分布定义满条件概率分布，依次对满条件概率分布进行抽样，得到样本的序列。

可以证明这样的抽样过程是在一个马尔可夫链上的随机游走，每一个样本对应着马尔可夫链的状态，平稳分布就是目标的联合分布。

整体成为一个马尔可夫链蒙特卡罗法，燃烧期之后的样本就是联合分布的随机样本。

**对比**

单分量Metropolis-Hastings算法

抽样会在样本点之间移动，但其间可能在某一些样本点上停留（由于抽样被拒绝）。

适合于满条件概率分布不容易抽样的情况，使用容易抽样的条件分作建议分布。

吉布斯抽样算法。

抽样会在样本点之间持续移动。

适合于满条件概率分布容易抽样的情况。