统计学习读书报告

09017244 郑健雄

1. 自己提出的问题的理解：
2. 吉布斯抽样比起直接使用Metropolis-Hastings算法的优势是什么？

讨论后的理解：

吉布斯抽样是Metropolis-Hastings算法的特例，而且是一种非常常用的特例。吉布斯抽样不需要考虑是否需要接受，并且使用了满概率分布，属于在MH算法里性质较好的一种情况，效率会高一点。

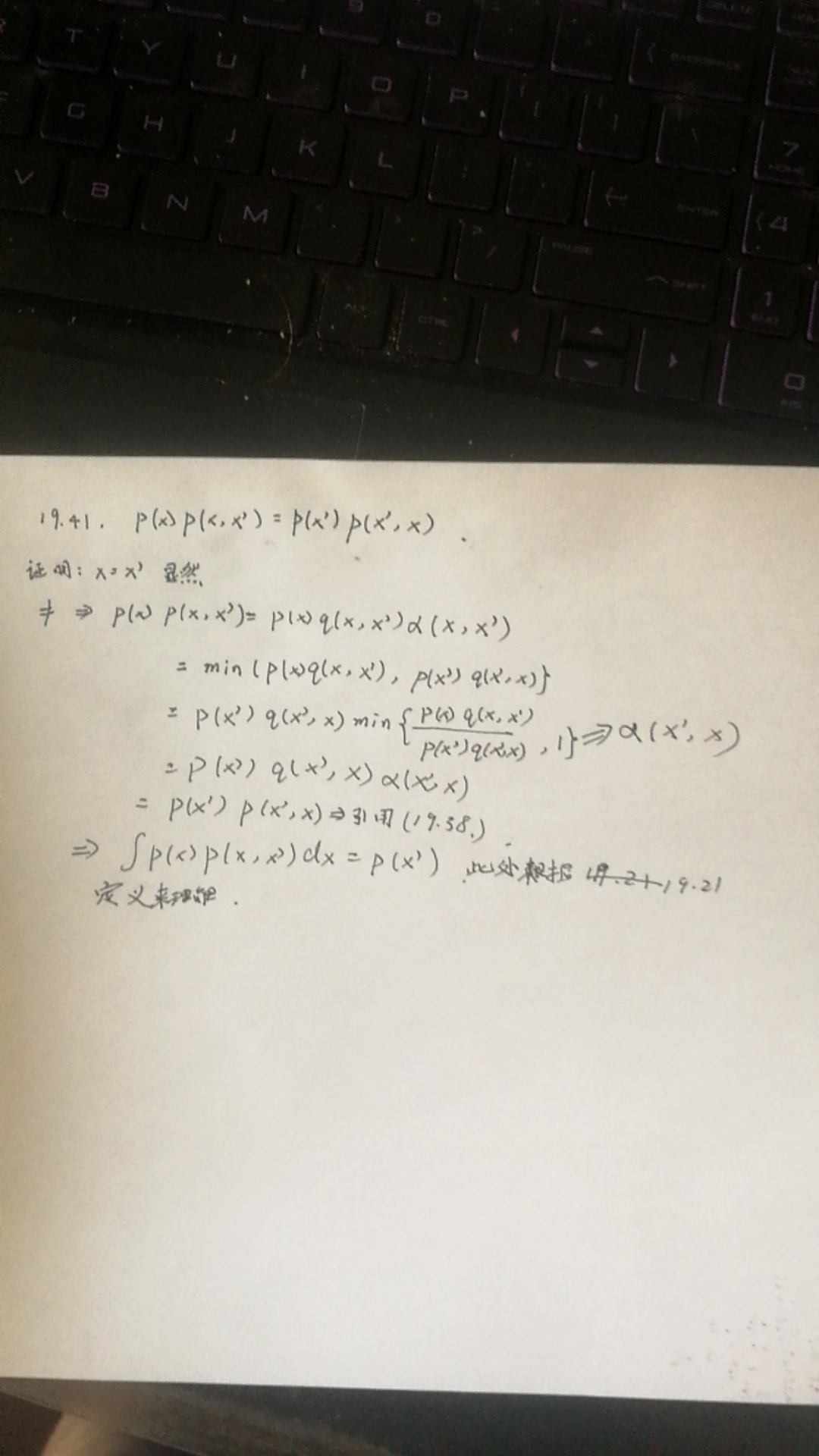
1. 如何理解p371的定理19.6中p(x)是平稳分布？

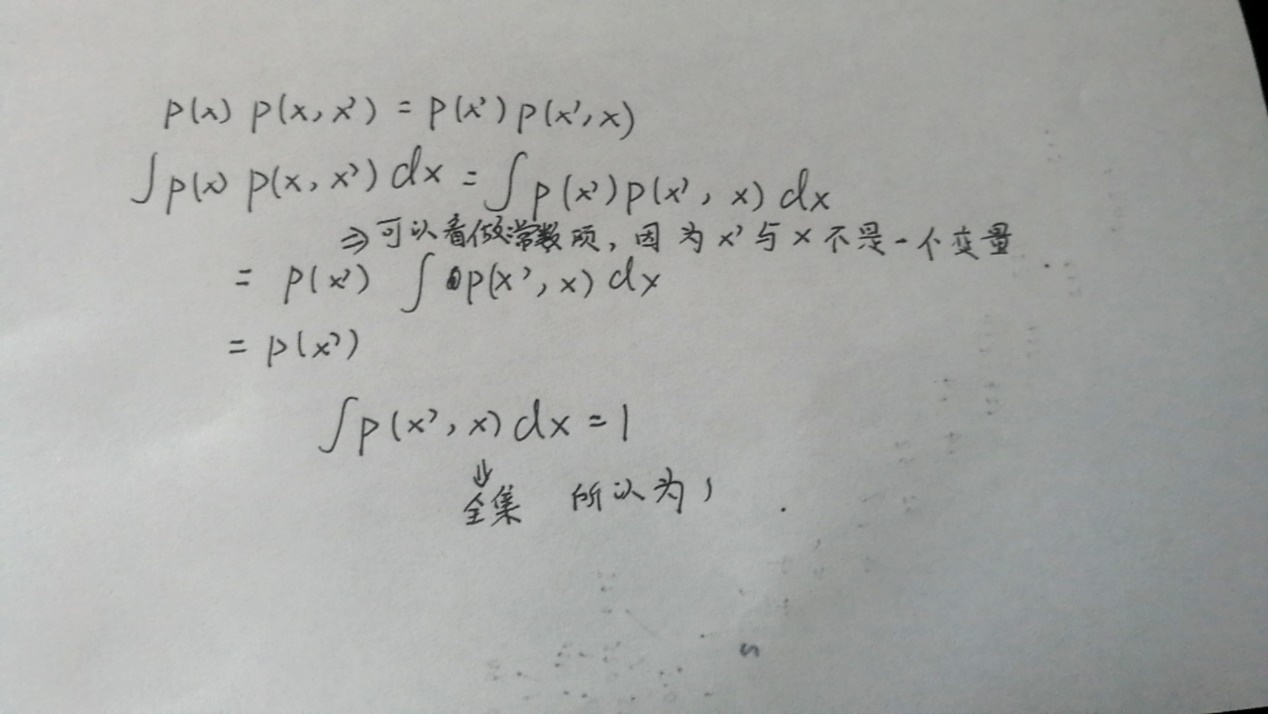
讨论后的理解：

此处的平稳分布可以在连续情况下思考，并不是说类似离散的情况，也就是稳定不动，在连续情况中，该公式是满足平衡方程的，从这个角度来看，可以看出这些状态都是可以互相抵达并且每次转移的概率是相等的的，也就是可逆的性质，符合平稳分布的定义。

1. 别人提出的问题的理解
2. P371 19.6推导下面的积分的推导和上面推导的最后一个等式。

自己的理解:

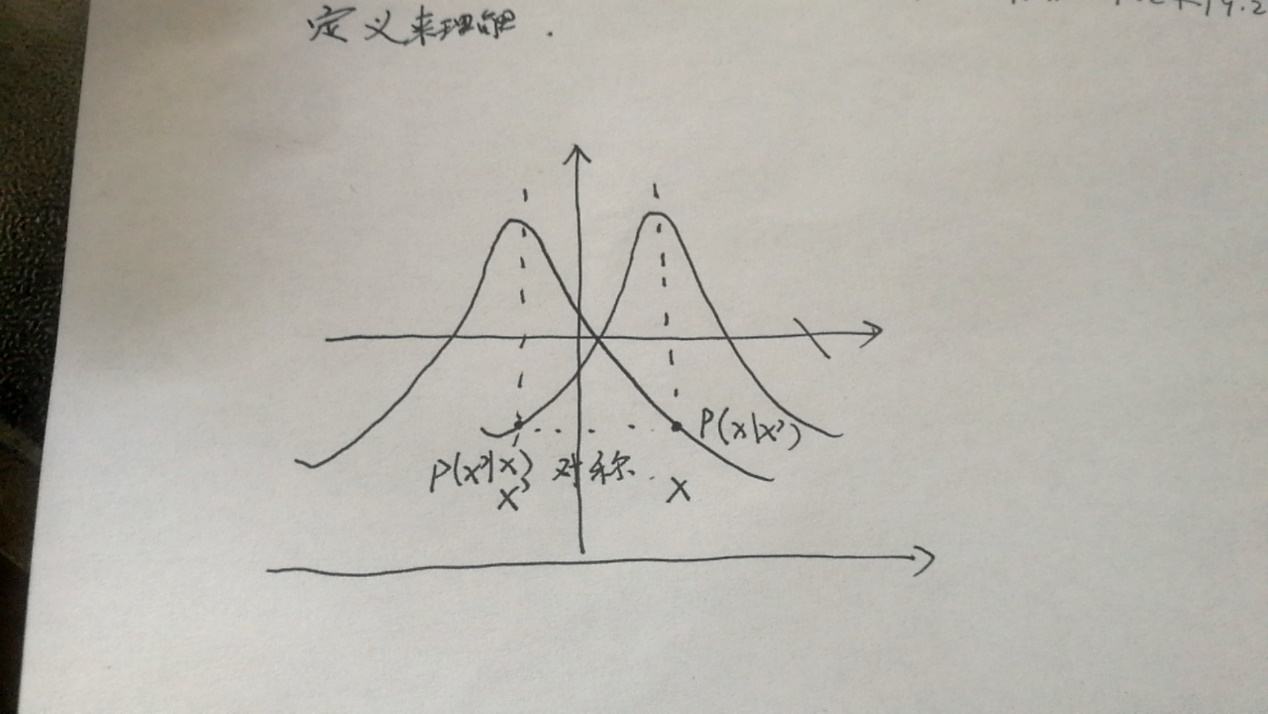




此处最后一个部分的理解是当从一个状态转移到所有状态的概率肯定为1，可以消除积分，从而证明等式。

1. P372建议分布第一种形式的第一个特例中，q(x,x')=p(x'|x)，而q(x',x)=p(x|x'),感觉两者好像不相等，但书上为什么说q(x,x')=q(x',x)

自己的理解：



此处给出的对条件概率的定义是多元正态分布，并且其均值为作为条件的状态的值，说明曲线的形状都是一致的，只有均值也就是对称性不同。以一元分布举例，根据概率本身在曲线上分布的位置以及几何图形的性质可以观察到p(x’|x)=p(x|x’)。

1. 如何理解吉布斯抽样适合于满条件概率分布容易抽样的情况，单分量Metropolis-Hasting算法适合于满条件概率分布不容易抽样的情况？

自己的理解：

吉布斯抽样选择的建议分布就是满概率分布，所以满概率分布容易抽样也就更加方便，而普通情况下不一定可以抽样满概率分布，所以只能用单分量的MH算法。

1. 怎么理解P375的“由于建议分布可能不被接收，MH算法可能在一些相邻的时刻不产生移动”？

自己的理解：

在使用该算法时候还需要计算接受概率,也就是用01分布抽样一个值,根据其是否大于接受分布值,大于则接受,小于则不接受,不接受的情况也就不产生移动。其思路与之前的接受-拒绝抽样类似。

1. 读书计划

1、本周完成的内容章节：复习19章，阅读完Speech and Language Processing第五章。

2、下周计划：复习20章，阅读部分第六章内容。

四、读书摘要及理解

1. 马尔可夫链蒙特卡罗法：该方法也属于蒙特卡罗法的一种，适合随机变量多元，密度函数不标准并且随机变量各个分量不独立的情况，其恰好符合马尔可夫链的特征。其思路是用平稳分布获取目标分布。为了满足条件，如何定义转移概率十分重要，一般使用可逆矩阵的情况来构建。常用的方法有MH算法及其特例吉布斯抽样。

2. MH算法：其使用的的转移核函数类似于接受-拒绝法，有建议分布和接受分布，该定义下的马尔可夫链是可逆的，平稳分布就是p(x)。一般选择的建议分布与接受分布与p(x)有着紧密的联系，比如条件概率等。有的时候如果满条件概率可以使用，可以用其加快计算。而使用单分量MH算法可以更方便的进行抽样，一般这种情况下多变量抽样很困难，但是单分量MH算法可能会造成更大的误差。

3. 吉布斯抽样：该算法是MH的特例，融合了其中比较好的性质，比如单分量抽样的思想并且无需考虑是否接受，效率较好，得到了广泛的应用。

思考：

抽样在统计学中是非常重要的问题，其可以帮助人们研究概率的内在规律。因此，如何对复杂的情况进行抽样将决定能否得到足够的数据以及能否验证模型的可靠性等问题。比较好的抽样方法其实是直接抽样，其十分直观并且有理论支持，而比较复杂的抽样方法一般都会遇到各种各样的问题，比如接受-拒绝法效率较差，马尔可夫链蒙特卡罗法生成下一个序列在很多情况中也是很困难的事情。尽管单分量和吉布斯抽样使用单分量方法避开了这个问题，但是其抽样结果的质量也会受到影响。