MCMC(三)MCMC采样和M-H采样

在MCMC(二)马尔科夫链中我们讲到给定一个概率平稳分布π, 很难直接找到对应的马尔科夫链状态转移矩阵P。而只要解决这个问题，我们就可以找到一种通用的概率分布采样方法，进而用于蒙特卡罗模拟。本篇我们就讨论解决这个问题的办法：MCMC采样和它的易用版M-H采样。

1. 马尔科夫链的细致平稳条件

在解决从平稳分布π, 找到对应的马尔科夫链状态转移矩阵P之前，我们还需要先看看马尔科夫链的细致平稳条件。定义如下：

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：

$$\pi(i)P(i,j) = \pi(j)P(j,i)$$

π(i)P(i,j)=π(j)P(j,i)

则称概率分布$\pi(x)$是状态转移矩阵P的平稳分布。

证明很简单,由细致平稳条件有：

$$\sum_{i=1}^{\infty}\pi(i)P(i,j) = \sum_{i=1}^{\infty}\pi(j)P(j,i) = \pi(j)\sum_{i=1}^{\infty}P(j,i) = \pi(j)$$

∑i=1∞π(i)P(i,j)=∑i=1∞π(j)P(j,i)=π(j)∑i=1∞P(j,i)=π(j)

将上式用矩阵表示即为：

$$\pi P = \pi$$

πP=π

即满足马尔可夫链的收敛性质。也就是说，只要我们找到了可以使概率分布$\pi(x)$满足细致平稳分布的矩阵P即可。这给了我们寻找从平稳分布π, 找到对应的马尔科夫链状态转移矩阵P的新思路。

不过不幸的是，仅仅从细致平稳条件还是很难找到合适的矩阵P。比如我们的目标平稳分布是$\pi(x)$,随机找一个马尔科夫链状态转移矩阵Q,它是很难满足细致平稳条件的，即：

$$\pi(i)Q(i,j) \neq \pi(j)Q(j,i)$$

π(i)Q(i,j)≠π(j)Q(j,i)

那么如何使这个等式满足呢？下面我们来看MCMC采样如何解决这个问题。

2. MCMC采样

由于一般情况下，目标平稳分布$\pi(x)$和某一个马尔科夫链状态转移矩阵Q不满足细致平稳条件，即

$$\pi(i)Q(i,j) \neq \pi(j)Q(j,i)$$

我们可以对上式做一个改造，使细致平稳条件成立。方法是引入一个$\alpha(i,j)$,使上式可以取等号，即：

$$\pi(i)Q(i,j)\alpha(i,j) = \pi(j)Q(j,i)\alpha(j,i)$$

问题是什么样的$\alpha(i,j)$可以使等式成立呢？其实很简单，只要满足下两式即可：

$$\alpha(i,j) = \pi(j)Q(j,i)$$

$$\alpha(j,i) = \pi(i)Q(i,j)$$

这样，我们就得到了我们的分布$\pi(x)$对应的马尔科夫链状态转移矩阵P，满足：

$$P(i,j) = Q(i,j)\alpha(i,j)$$

也就是说，我们的目标矩阵P可以通过任意一个马尔科夫链状态转移矩阵Q乘以$\alpha(i,j)$得到。$\alpha(i,j)$我们有一般称之为接受率。取值在[0,1]之间，可以理解为一个概率值。即目标矩阵P可以通过任意一个马尔科夫链状态转移矩阵Q以一定的接受率获得。这个很像我们在MCMC(一)蒙特卡罗方法第4节讲到的接受-拒绝采样，那里是以一个常用分布通过一定的接受-拒绝概率得到一个非常见分布，这里是以一个常见的马尔科夫链状态转移矩阵Q通过一定的接受-拒绝概率得到目标转移矩阵P,两者的解决问题思路是类似的。

好了，现在我们来总结下MCMC的采样过程。

1）输入我们任意选定的马尔科夫链状态转移矩阵Q，平稳分布$\pi(x)$，设定状态转移次数阈值n_1，需要的样本个数n_2

2）从任意简单概率分布采样得到初始状态值x_0

3）for $t = 0$ to $n_1 + n_2 - 1$:

a) 从条件概率分布$Q(x|x_t)$中采样得到样本x_*

b) 从均匀分布采样$u \sim uniform[0,1]$

c) 如果$u < \alpha(x_t, x_*) = \pi(x_*)Q(x_*, x_t)$, 则接受转移$x_t \rightarrow x_*$，即$x_{t+1} = x_*$

d) 否则不接受转移，$t = max(t-1, 0)$

样本集$(x_{n_1}, x_{n_1+1}, ..., x_{n_1+n_2-1})$ (xn1,xn1+1,...,xn1+n2−1)即为我们需要的平稳分布对应的样本集。

上面这个过程基本上就是MCMC采样的完整采样理论了，但是这个采样算法还是比较难在实际中应用，为什么呢？问题在上面第三步的c步骤，接受率这儿。由于$\alpha(x_t, x_*)$ α(xt,x∗)可能非常的小，比如0.1，导致我们大部分的采样值都被拒绝转移，采样效率很低。有可能我们采样了上百万次马尔可夫链还没有收敛，也就是上面这个n_1 n1要非常非常的大，这让人难以接受，怎么办呢？这时就轮到我们的M-H采样出场了。

3. M-H采样

M-H采样是Metropolis-Hastings采样的简称，这个算法首先由Metropolis提出，被Hastings改进，因此被称之为Metropolis-Hastings采样或M-H采样

M-H采样解决了我们上一节MCMC采样接受率过低的问题。

我们回到MCMC采样的细致平稳条件：

$$\pi(i)Q(i,j)\alpha(i,j) = \pi(j)Q(j,i)\alpha(j,i)$$

π(i)Q(i,j)α(i,j)=π(j)Q(j,i)α(j,i)

我们采样效率低的原因是$\alpha(i,j)$ α(i,j)太小了，比如为0.1，而$\alpha(j,i)$ α(j,i)为0.2。即：

$$\pi(i)Q(i,j) \times 0.1 = \pi(j)Q(j,i) \times 0.2$$

π(i)Q(i,j)×0.1=π(j)Q(j,i)×0.2

这时我们可以看到，如果两边同时扩大五倍，接受率提高到了0.5，但是细致平稳条件却仍然是满足的，即：

$$\pi(i)Q(i,j) \times 0.5 = \pi(j)Q(j,i) \times 1$$

π(i)Q(i,j)×0.5=π(j)Q(j,i)×1

这样我们的接受率可以做如下改进，即：

$$\alpha(i,j) = min\{\frac{\pi(j)Q(j,i)}{\pi(i)Q(i,j)}, 1\}$$

α(i,j)=min{π(j)Q(j,i)π(i)Q(i,j),1}

通过这个微小的改造，我们就得到了可以在实际应用中使用的M-H采样算法过程如下：

1）输入我们任意选定的马尔科夫链状态转移矩阵Q Q，平稳分布$\pi(x)$ π(x)，设定状态转移次数阈值n_1 n1，需要的样本个数n_2 n2

2）从任意简单概率分布采样得到初始状态值x_0 x0

3）for $t = 0$ t=0 to $n_1 + n_2 - 1$ n1+n2−1:

a) 从条件概率分布$Q(x|x_t)$ Q(x|xt)中采样得到样本x_* x∗

b) 从均匀分布采样$u \sim uniform[0,1]$ u∼uniform[0,1]

c) 如果$u < \alpha(x_t, x_*) = \min\{\frac{\pi(j)Q(j,i)}{\pi(i)Q(i,j)}, 1\}$ u<α(xt,x∗)=min{π(j)Q(j,i)π(i)Q(i,j),1} , 则接受转移$x_t \rightarrow x_*$ xt→x∗ ，即$x_{t+1} = x_*$ xt+1=x∗

d) 否则不接受转移，$t = \max(t-1, 0)$t=max(t−1,0)

样本集$(x_{n_1}, x_{n_1+1}, ..., x_{n_1+n_2-1})$ (xn1,xn1+1,...,xn1+n2−1)即为我们需要的平稳分布对应的样本集。

很多时候，我们选择的马尔科夫链状态转移矩阵QQ如果是对称的，即满足$Q(i,j) = Q(j,i)$ Q(i,j)=Q(j,i),这时我们的接受率可以进一步简化为：

$$\alpha(i,j) = \min\{\frac{\pi(j)}{\pi(i)}, 1\}$$

α(i,j)=min{π(j)π(i),1}

4. M-H采样实例

为了更容易理解，这里给出一个M-H采样的实例。

完整代码参见我的github: https://github.com/ljpzzz/machinelearning/blob/master/mathematics/mcmc_3_4.ipynb

在例子里，我们的目标平稳分布是一个均值3，标准差2的正态分布，而选择的马尔可夫链状态转移矩阵$Q(i,j)$ Q(i,j)的条件转移概率是以i i为均值,方差1的正态分布在位置j j的值。这个例子仅仅用来让大家加深对M-H采样过程的理解。毕竟一个普通的一维正态分布用不着去用M-H采样来获得样本。

代码如下：

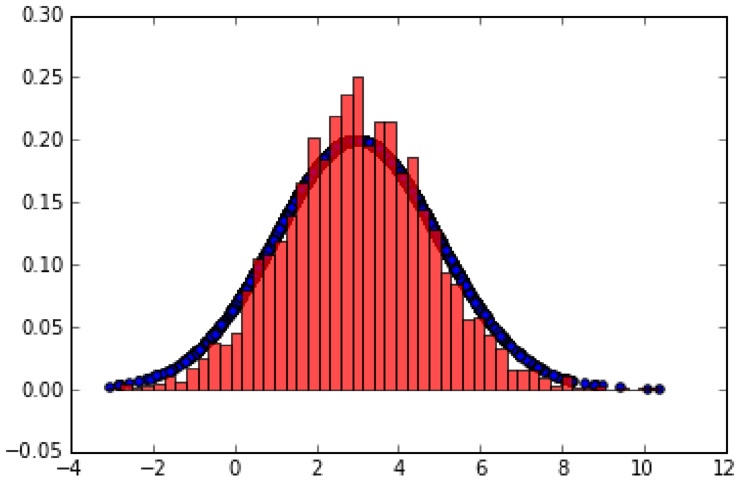
复制代码

```
import random
import math
from scipy.stats import norm
import matplotlib.pyplot as plt
%matplotlib inline

def norm_dist_prob(theta):
    y = norm.pdf(theta, loc=3, scale=2)
    return y

T = 5000
pi = [0 for i in range(T)]
sigma = 1
t = 0
while t < T-1:
    t = t + 1
    pi_star = norm.rvs(loc=pi[t - 1], scale=sigma, size=1, random_state=None)
    alpha = min(1, (norm_dist_prob(pi_star[0]) / norm_dist_prob(pi[t - 1])))

    u = random.uniform(0, 1)
    if u < alpha:
        pi[t] = pi_star[0]
    else:
        pi[t] = pi[t - 1]
```

```
plt.scatter(pi, norm.pdf(pi, loc=3, scale=2))
num_bins = 50
plt.hist(pi, num_bins, normed=1, facecolor='red', alpha=0.7)
plt.show()
```

复制代码

输出的图中可以看到采样值的分布与真实的分布之间的关系如下，采样集还是比较拟合对应分布的。



5. M-H采样总结

M-H采样完整解决了使用蒙特卡罗方法需要的任意概率分布样本集的问题，因此在实际生产环境得到了广泛的应用。

但是在大数据时代，M-H采样面临着两大难题：

1）我们的数据特征非常的多，M-H采样由于接受率计算式$\frac{\pi(j)Q(j,i)}{\pi(i)Q(i,j)}$ π(j)Q(j,i)π(i)Q(i,j)的存在，在高维时需要的计算时间非常的可观，算法效率很低。同时$\alpha(i,j)$ α(i,j)一般小于1，有时候辛苦计算出来却被拒绝了。能不能做到不拒绝转移呢？

2）由于特征维度大，很多时候我们甚至很难求出目标的各特征维度联合分布，但是可以方便求出各个特征之间的条件概率分布。这时候我们能不能只有各维度之间条件概率分布的情况下方便的采样呢？

Gibbs采样解决了上面两个问题，因此在大数据时代，MCMC采样基本是Gibbs采样的天下，下一篇我们就来讨论Gibbs采样。

（欢迎转载，转载请注明出处。欢迎沟通交流： liujianping-ok@163.com）

分类: 0040. 数学统计学
标签: 机器学习中的数学
好文要顶 关注我 收藏该文
刘建平Pinard
关注 - 14

« 上一篇：MCMC(二)马尔科夫链
» 下一篇：MCMC(四)Gibbs采样

评论列表

#1楼 2017-08-03 18:20 pmonkeychen
您好，请问第二第三章在叙述算法过程时，3)中的a)里应该都是从Q中生成样本而不是P吧?
支持(0)反对(0)

#2楼[楼主] 2017-08-03 18:24 刘建平Pinard
@ pmonkeychen
你好，这儿的确写错了，已改正。感谢指出错误。
支持(0)反对(0)

#3楼 2017-09-10 18:22 DTStor
你好，请问代码
pi_star = norm.rvs(loc=theta[t - 1], scale=sigma, size=1, random_state=None)
中的theta在哪儿定义的？谢谢
支持(0)反对(0)

#4楼 2017-09-10 19:20 DTStor
知道了，
pi_star = norm.rvs(loc=theta[t - 1], scale=sigma, size=1, random_state=None)
中的theta应该是pi
支持(0)反对(0)

#5楼[楼主] 2017-09-11 10:30 刘建平Pinard
@ DTStor
你好，这里的确如你所说，应该是pi,已修改，感谢指出错误。
支持(0)反对(0)

#6楼 2017-10-03 00:00 dethrive
请问，第2和第三部分的算法中“d) 否则不接受转移，t = t - 1”，如果 t = 0 时候，算法就直接拒绝了，那么 t = -1，后面的算法还能进行下去吗？或者说这里的 t 是不是要理解为第 t 次有效的样本采集啊?
支持(0)反对(0)

#7楼 2017-10-03 00:16 dethrive
@ 刘建平Pinard
还有应该是均值为3，标准差为2
支持(0)反对(0)

#8楼[楼主] 2017-10-09 11:06 刘建平Pinard
@ dethrive
你好，这里我写的的确有点问题。改成了这个样子，避免初始拒绝的问题。

$t = \max(t-1, 0)$t=max(t−1,0)
那个均值和标准差的笔误也改了。
感谢指出这么多的错误。
支持(0)反对(0)

#9楼 2017-10-31 10:42 王小熊
你好，我想请问程序中在求alpha的时候，为什么没有乘以pi，公式里是有的
支持(0)反对(0)

#10楼[楼主] 2017-10-31 11:23 刘建平Pinard
@ 王小熊
你好，这里采用的是这个公式计算的αα

$$\alpha(i,j) = \min\{\frac{\pi(j)}{\pi(i)}, 1\}$$

α(i,j)=min{π(j)π(i),1}

即假定了$Q(i,j) = Q(j,i)$ Q(i,j)=Q(j,i)

当然我上面的例子里面Q并不对称。这个例子实际使用的是最早的Metropolis采样的公式，这个公式并不管Q对称与否，直接用了上式。

你可以修改下我的程序，按照我上面的M-H采样公式更新αα，再画图看看结果。：）
支持(0)反对(0)

#11楼 2017-10-31 11:24 王小熊
@ 刘建平Pinard
谢谢！
支持(0)反对(0)

#12楼 2017-12-12 16:34 爱吃香菜的JQ
你好，算法 3）：d）否则不接受转移，t=max(t−1,0) 这个是拒绝本次采样，那程序中为什么是else: pi[t] = pi[t - 1]?
支持(1)反对(0)

#13楼[楼主] 2017-12-12 17:17 刘建平Pinard
@ 爱吃香菜的JQ
你好，这里不接受转移的话，t就是t-1，取max(t−1,0)是为了避免在t=0的时候的t变成负数。
不接受转移的话，t就是t-1，那么pi[t] = pi[t - 1]
支持(0)反对(0)

#14楼 2017-12-12 18:03 爱吃香菜的JQ
@ 刘建平Pinard
我认为t=t-1不等于pi[t] = pi[t - 1]。t=t-1意味者拒绝这次采样，pi[t] = pi[t - 1]意味着 连续两个pi是同一个值，就如下：
序号 类型 数值
23 float64 4.63203414485
24 float64 4.63203414485
支持(0)反对(0)

#15楼[楼主] 2017-12-13 10:48 刘建平Pinard

@ 爱吃香菜的JQ
你好，这里你说的的确有道理，这里算法和下面的例子上逻辑稍有不同。

在算法里，维度目标是得到n_1n1个真正的转移后，后面的n_2n2个样本就是我们的需要的数据，所以我使用了t=t-1, 这样被拒绝的样本就不进入采样样本序列。

而在下面的例子里面，就算不是真正的转移（被决绝），我们也把它放入了样本序列，所以有些mismatch。

不过这个不同对算法的理解影响并不大。
支持(0)反对(0)

#16楼 2018-02-10 23:31 扼杀
α(i,j)α(i,j)我们有一般称之为接受率。取值在[0,1]之间，可以理解为一个概率值。 π(j)*Q(j,i)这个可以得到一个数，而不是矩阵吗?
支持(0)反对(0)

#17楼[楼主] 2018-02-11 13:50 刘建平Pinard
@ 扼杀
你好，QQ是一个矩阵，而$Q(j,i)$ Q(j,i)是矩阵对应位置的一个标量，所以$\pi(j)*Q(j,i)$ π(j)∗Q(j,i)可以得到一个数，而$\pi(j)*Q$ π(j)∗Q可以得到一个一个矩阵。
支持(2)反对(0)

#18楼 2018-02-22 22:41 扼杀
@ 刘建平Pinard
谢谢博主大神
支持(0)反对(0)

#19楼 2018-03-12 21:26 泡泡糖nana
谢谢博主大神，对我很有帮助!
支持(0)反对(0)

#20楼 2018-04-30 09:14 kylin0228
博主您好，想请问一下细致平稳条件一般不成立是不是说：这是pai(x)的充分不必要条件? ？就是说满足这个条件一定是平稳分布了，但是不满足也不代表不是?
支持(0)反对(0)

#21楼[楼主] 2018-05-01 00:01 刘建平Pinard
@ kylin0228
我说的应该不是这个意思。
满足细致平稳条件，那么这作为一个充分条件，可以推导出$\pi(x)$π(x)。

细致平稳条件一般不成立，这是因为这个矩阵很难直接找出来。也就是说，我随便给你一个概率分布，比如[0.3,0.5,0.2](只是举个例子)，你很难给出它所对应的满足细致平稳条件的转移矩阵。
支持(0)反对(0)

#22楼 2018-05-01 08:35 kylin0228
博主您好
后来我查阅了一些资料
细致平稳条件是一个充分不必要条件

即不满足细致平稳条件不一定不是平稳分布
参考链接：https://www.zhihu.com/question/37000071/answer/131542485
支持(0)反对(0)

#23楼[楼主] 2018-05-01 23:09 刘建平Pinard
@ kylin0228
你好，细致平稳条件是一个充分不必要条件，我21楼也说了"这作为一个充分条件"。

只是"细致平稳条件一般不成立",不是因为它是一个充分不必要条件。而是这样的矩阵难以直接找出。
支持(0)反对(0)

#24楼 2018-05-02 11:28 kylin0228
嗯嗯 get到博主的意思了
谢谢博主大神!
支持(0)反对(0)

#25楼 2018-05-19 09:34 蜉蝣2015
@ 刘建平Pinard
从您的代码上来看，我是这样理解的 1）存储状态序列里面的值均为接收值，只是有发生重复的时候
2）这里这个序列没有体现出来 超过 n1阈值的地方。
支持(0)反对(0)

#26楼[楼主] 2018-05-19 12:09 刘建平Pinard
@ 蜉蝣2015
1）是这样的。
2） 在实际工程应用中，我们一般不判断阈值，都是默认采样到某一个数量，后面认为就是我们需要的样本。如果计算量大的时候，这个“数量阈值”有时候都不管了。一般问题不大。
支持(1)反对(0)

#27楼 2018-05-19 19:37 蜉蝣2015
还是有些不明白。1）根据文中所述，MCMC和MH两种方法中的detail balance condition，对所有的用分布 Q生成的样本点都成立，这点您同意么?
2）为什么通过 比较 u和接收率a的值，就能确定Q 生成的样本点，也可以作为P的样本点? 除了从rejection sampling 方法中，感性的比较去理解，有啥具体的证明或者定量的分析不?
支持(0)反对(0)

#28楼[楼主] 2018-05-19 22:53 刘建平Pinard
@ 蜉蝣2015
你好，1） 是的，通过一定的拒绝率，我们用Q模拟得到了满足细致平稳条件的分布P。
2） 如果你对MCMC进一步的数学原理和演进感兴趣，推荐你读这篇文章:
http://stat.wharton.upenn.edu/~stjensen/stat542/lecture14.mcmchistory.pdf
支持(1)反对(0)

#29楼 2018-05-20 09:08 蜉蝣2015
@ 刘建平Pinard
好的 谢谢。
支持(0)反对(0)

#30楼 2018-06-22 11:40 忆梦涟

请问，之前的MCMC和M-H算法是基于联合概率吗，看起来中间用到的都是条件概率，有些不明白，谢谢！
支持(0)反对(0)

#31楼[楼主] 2018-06-22 23:27 刘建平Pinard
@ 忆梦涟
你好，这里是条件概率。
你可以理解为有一个概率分布函数P(A)，当前有一个点在分布的位置A1，下一个点采样得到分布位置A2的概率。也就是P(A2|A1)。这个P(A2|A1)是分布P(A)中对应的两点之间的条件概率。你可以理解为是一个状态转移概率。
支持(1)反对(0)

#32楼 2018-06-29 16:34 renminghuang
老师，关于细致平稳条件，有个地方不懂。
pi(i)是随机变量状态为i时的概率，P(i,j)是随机变量从状态i转移至状态j的概率，它是一个条件概率，那么两者相乘即随机变量在状态i和状态j时的联合分布概率，那么细致平稳条件等式两侧应该是恒相等的，为什么会大多数情况下不相等，而是需要再乘以个接受率呢？
谢谢。

编辑：
是不是因为状态存有时序的原因？即时刻t状态i时刻t+1状态j的联合概率分布不一定等于时刻t状态j时刻t+1状态i的联合概率分布。
支持(0)反对(0)

#33楼[楼主] 2018-07-01 22:04 刘建平Pinard
@ renminghuang
你好，这里的原因应该是这个P矩阵很难寻找，我们使用的是另外一个Q矩阵，这个Q矩阵和P矩阵是不一样的，所以需要一个接受率来用Q慢慢拟合P。

如果我们直接拿到了P矩阵，那么的确就是恒等了。
支持(0)反对(0)

#34楼 2018-07-07 16:02 潇洒走着
有两个疑问希望博主能解答下：
1. 为什么要设置u来跟α比较，判断是否接受。
2. 代码中，在上一个的采样值作为均值，1为方差的高斯分布上做采用， 为什么用上一个采样的样本做当前采样的高斯分布均值？
支持(0)反对(0)

#35楼[楼主] 2018-07-08 22:25 刘建平Pinard
@ 潇洒走着
你好，第一个问题其实我第二节已经讲过了。你可以看看，简言之就是Q不是我们想要的P，通过一定的接收率来让Q近似于P。

第二个问题，这里其实是我们假设的，上面我也写了，是： “选择的马尔可夫链状态转移矩阵Q(i,j)的条件转移概率是以i为均值,方差1的正态分布在位置j的值”
支持(1)反对(0)

#36楼 2018-07-26 10:49 华师大小公主

@ 刘建平Pinard
博主您好，MH算法中，平稳分布π没有更新，转移概率q(i,j)也没有更新，相应的α(i,j)也不会改变，而且从开始采样的时候，由于加入了α就满足平稳分布了，为什么还需要采样到一定阈值之后的样本，才算是分布的真实样本？不应该从t=1次采样开始就是分布的真实样本吗？
支持(0)反对(0)

#37楼[楼主] 2018-07-26 11:02 刘建平Pinard
@ 华师大小公主
你好！加入了接受率后，的确可以使Q近似于P了。但是如果你看了上一届马尔科夫链后，你会发现，就算拿到了P，你也得采样到一定的数目，让采样分布收敛后才能得到你想要的样本。
支持(0)反对(0)

#38楼 2018-07-26 11:18 华师大小公主
博主，我去看了上节的，第三小节您确实提到采样n次后的样本才是平稳分布的对应样本集，但是似乎没有解释原因（可能没理解到）。

所以我还是有点不明白，为什么需要采样到一定数目后才是想要的？因为并没有看到哪个分布在更新，那么所谓采样分布收敛，是如何收敛的？

您指的采样分布是指π呢还是转移概率q？如果是π，在上面MH的算法中，π作为输入，是已知的；如果是q，算法中没有看到q的更新？所以，何来的采样分布收敛呢？
支持(0)反对(0)

#39楼[楼主] 2018-07-27 10:40 刘建平Pinard
@ 华师大小公主
你好！ 原因你可以看我上一篇的python代码，也可以跑一遍。采样n次后，我们的分布ππ就稳定了，这个稳定分布的ππ就是我们想要采样的分布。

具体收敛我没有证明，不过通过一个实际的例子和代码展示了。

采样的分布是ππ，这个ππ在乘以转换矩阵P后分布会变，当ππ稳定后就是采样分布收敛
支持(0)反对(0)

#40楼 2018-07-27 10:59 华师大小公主
@ 刘建平Pinard
π是平稳分布，按照定义πp=π，为什么π乘以p后π还会变呢？
支持(0)反对(0)

#41楼 2018-07-27 11:19 华师大小公主
@ 刘建平Pinard
看完LDA，我在想，是不是因为每次采样完，条件概率分布Q(x|xt)发生了改变，所以需要等到Q(x|xt)收敛后，采样才算有效？
支持(0)反对(0)

#42楼[楼主] 2018-07-27 11:35 刘建平Pinard
@ 华师大小公主
你好，最开始你采样使用的初始ππ是任意选择的，不是我们的平稳分布ππ。所以乘了以后会变。
条件概率分布Q(x|xt)的分布函数是不变的，但是每次$X_t$xt会变，这样采样得到的X x也是会变的。不过根本原因还是你最开始使用的ππ不是平稳分布，所以采样无效。
支持(0)反对(0)

#43楼 2018-09-18 14:13 南鱼知刃
@ 刘建平Pinard
@ 华师大小公主
" 你好！加入了接受率后，的确可以使Q近似于P了。但是如果你看了上一届马尔科夫链后，你会发现，就算拿到了P，你也得采样到一定的数目，让采样分布收敛后才能得到你想要的样本。"

您好，上面是您回复这位网友的一段话，个人认为上一节马尔科夫链中的P一开始是没有收敛的（即一开始不满足细致平稳条件），所以需要采样多次（不过在上一节最后MH的算法描述中确实没看到P更新，很奇怪）；但这一节中的P一开始就是满足细致平稳条件的（只是不能直接得到P，而是用Q通过拒绝-接受来模拟P而已），所以在MCMC采样中应该一开始的采样样本就能用啊；
以上两点疑问，希望您有空的时候帮忙解答下，感激不尽！
支持(0)反对(0)

#44楼[楼主] 2018-09-18 14:33 刘建平Pinard
@ 南鱼知刃
你好！
首先，马尔科夫状态转移矩阵没有收敛一说，只有采样的概率分布有收敛一说。对应上一篇中的符号，就是采样的概率分布开始是$\pi(i)$ π(i),随着采样的进行，收敛到我们需要的采样分布$\pi(j)$ π(j),即：

$$\pi(j) = \sum_{i=0}^{\infty} \pi(i)P_{ij}$$

π(j)=∑i=0∞π(i)Pij

采样多次的原因是希望$\pi(i)$ π(i)慢慢向$\pi(j)$ π(j)靠拢。
支持(0)反对(0)

#45楼 2018-09-27 11:22 南鱼知刃
@ 刘建平Pinard
您好，在MCMC采样中，既然已知了任意转移矩阵Q和和平稳分别$\pi(x)$ π(x)，就可以计算得到目标矩阵P，那在第三步中为什么不直接根据$P(x|x_t)$ P(x|xt)采样，而要根据$Q(x|x_t)$ Q(x|xt)采样？
支持(0)反对(0)

#46楼[楼主] 2018-09-28 10:26 刘建平Pinard
@ 南鱼知刃
你好，目标矩阵P不是计算得到的，而是采样$Q(x|x_t)$ Q(x|xt),并通过接受拒绝这样的方式间接得到的。没法直接采样$P(x|x_t)$ P(x|xt)
支持(0)反对(0)
刷新评论刷新页面返回顶部
注册用户登录后才能发表评论，请 登录 或 注册，访问网站首页。