MCMC的理解

MCMC统计模拟方法如果从纯公式的角度一开始并不好理解。我刚开始时的理解过程是这样的。

1、细致平稳条件要求：

（1）

这个容易理解。根据细致平稳条件，对于不满足条件的Markov转移矩阵进行改造，也好理解。

2、改造Markov转移矩阵：

(2)

(3)

这时我就碰到了问题：首先为什么要将称为接受概率，那么拒绝的概率给了谁？其次，根据公式(2)，我们可以知道，从而导致，从而导致细致平稳条件不满足。最后，为什么经典算法中不直接构造新的状态转移概率矩阵进行计算，而弄出一个接受和拒绝的过程。

其实三个问题的症结是一个，那就是一般介绍MCMC的资料中没有说明状态自我转移永远是满足细致平稳条件的，自己转给自己，当然转入和转出相等。因此在构造MCMC时，每个状态多出的未转移的概率都留给了q(i,i)。下面是我给出的直观解释：

首先细致平稳条件的直观含义如下图。

p(i)

p(j)

p(i)q(i,j)

p(j)q(j,i)

p(j)q(j,j)

p(i)q(i,i)

图1 细致平稳条件的状态转移概率图

从状态i转移到状态j的概率和从状态j转移回来的概率相同，从而达到了状态i和状态j概率的稳定。由于对于普通的Markov转移矩阵[q(i,j)]，状态i转移到状态j出去的概率不等于从状态j转回到状态i的概率。经典的构造方法是通过转移概率乘以，同时降低状态转移出去和转移回来的概率，使得二者相等，多出来的转移概率换成了自我转移，从而使得状态自我转移的概率大大加大，从而可能导致Markov链收敛速度大幅度下降。如下图：

p(i)

p(j)

p(i)q(i,j)p(j)q(j,i)

p(j)q(j,i)p(i)q(i,j)

p(i)(q(i,i)+q(i,j)(1-p(j)q(j,i)))

p(j)(q(j,j)+q(j,i)(1-p(i)q(i,j)))

图2 经典方法的状态概率转移图

以三状态来说明经典方法状态转移矩阵的变化。三状态平稳分布为[p(1) p(2) p(3)]，状态转移概率矩阵为：

(4)

采用经典方法后构成的状态转移矩阵为：

(5)

因此，Metropolis-Hastings方法，就是取i,j两状态之间最小转移概率作为二者之间转入和转出的概率，转出去概率大的状态增加了自我转移的概率。假设状态i转出到状态j的概率大，如下图：

p(i)

p(j)

p(j)q(j,i)

p(j)q(j,i)

p(i)(q(i,i)+q(i,j))-p(j)q(j,i)

p(j)q(j,j)

图3 Metropolis-Hastings方法的状态概率转移图

因此，如果要构造最快收敛的转移方法思路就是调整每个状态的自我转移概率，使它最接近稳态下的自我转移概率，多余的概率都要转移出去同时保证细致平稳条件。