1. EM算法

EM算法是一种迭代算法，每次迭代分为两步: ①求期望(E)②求极大(M)，所以这一算法又称做期望极大算法。

9.1 EM算法的引入

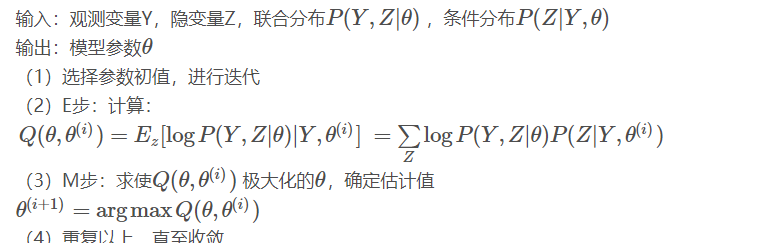
**三硬币模型：**给出抛掷硬币的结果，估算三硬币正面朝上的参数。

**完全数据**：观测随机变量的数据+隐随机变量

**不完全数据**：仅观测随机变量

9.1.1里 EM算法部分主要讲的就算极大似然估计，在我们已知样本所服从的概率分布模型和样本的情况下，估算出模型中的参数

EM算法的步骤如下：

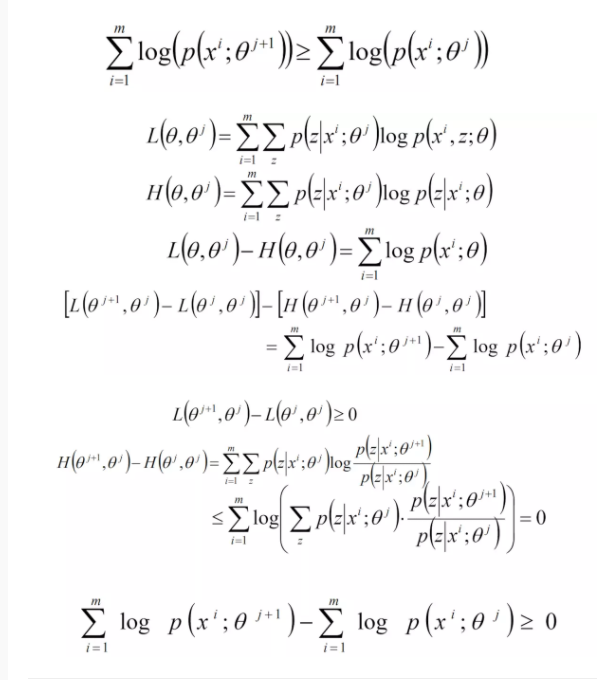


可以观察得到，其实EM算法是两个过程的交替：①是计算期望（E步），利用对隐藏变量的现有估计值，计算其最大似然估计值；在这里Q函数是EM算法的核心，即在E步所求得的最大似然估计值②是最大化（M），最大化在E步上求得的最大似然值来计算参数的值。M步上找到的参数估计值被用于下一个E步计算中，这个过程不断交替进行。

从前面的例子就可以算出来，每一步都在不断的更新参数，直到这一步的Q函数与上一步的Q函数满足一定的收敛条件为止。

9.2 EM算法的收敛性

EM算法的收敛性只要我们能够证明对数似然函数的值在迭代的过程中是增加的即可。

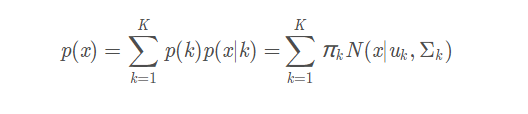


9.3 EM算法在高斯混合模型中的应用

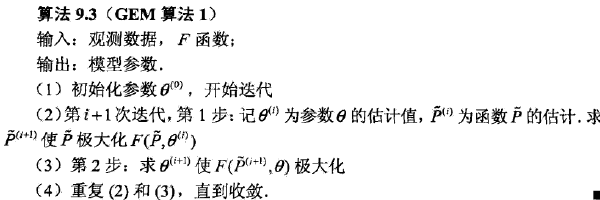
混合高斯模型的本质就是融合几个单高斯模型，来使得模型更加复杂，从而产生更复杂的样本。理论上，如果某个混合高斯模型融合的高斯模型个数足够多，它们之间的权重设定得足够合理，这个混合模型可以拟合任意分布的样本

它通过求解两个高斯模型，并通过一定的权重将两个高斯模型融合成一个模型，即最终的混合高斯模型。这个混合高斯模型可以产生这样的样本。

更一般化的描述为：假设混合高斯模型由K个高斯模型组成（即数据包含K个类），则GMM的概率密度函数如下：

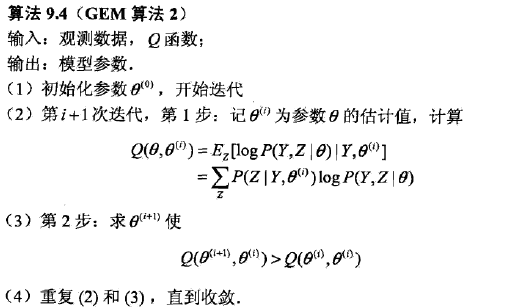


9.4 EM算法的推广

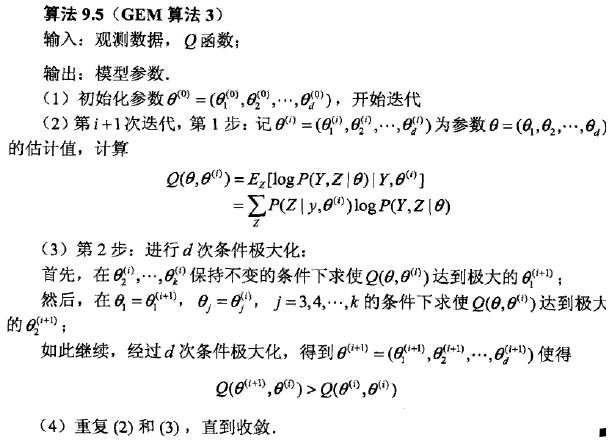


在GEM算法1中，有时求Q(theta，theta(i))的极大化是很困难的。

GEM算法2和GEM算法3并不是直接求theta(i+1)使Q达到极大的theta，而是找一个theta(i+1)使得Q(theta(i+1), theta(i)) >Q(theta(i), theta(i))



当参数theta的维数为d(d>=2)时，可采用一种特殊的GEM算法，它将EM算法的M步分解为d次条件极大化，每次只改变参数向量的一个分量，其余分量不改变。



GEM算法的特点是每次迭代增加F函数值(并不一定是极大化F函数)，从而增加似然函数值。