读书报告

71117203袁佳怡

1. 自己提出的问题的理解：
2. 提出的问题1：怎么理解EM算法不能保证找到全局最优解？

讨论后的理解：不同的初值得到的结果都是不一样的，只能够收敛到局部，所以对于初值不一样的，得到的局部最优解也是不相同的。

1. 别人提出的问题的理解（选择几个问题罗列，并给出理解）：
2. 问题2：GEM算法2和3可以达到GEM算法1一样的效果吗？

自己的理解：三种方法都是GEM的变形，只要每次都增加似然函数值并且收敛，那么它们的结果应该差不多。

1. 问题3：高斯混合模型可以用在什么地方？可以举一个具体的例子吗？

自己的理解：在图像处理中，当要区分目标和背景的时候，由于灰度分布图是不同的，需要混合建模来进行区分。

1. 读书计划

1、本周完成的内容章节：《统计学习方法》第九章

2、下周计划：《统计学习方法》第十章

四、读书摘要

1、在进行概率模型的参数估计时，如果变量全是可以直接观测的，就可以直接用极大似然估计或贝叶斯估计；如果模型中含有隐变量，此时可以用EM算法来进行估计。EM算法就是含有隐变量的概率模型参数的极大似然估计。EM算法通过迭代来进行参数估计，每次迭代分为E和M两步，E步——求期望（计算Q函数）；M步——求极大，求令Q函数极大的θ值并对其进行参数更新。

2、EM算法是一种非监督模型，是含有缺失数据的概率模型参数的极大似然估计法。不完全数据：观测数据X（观测随机变量得到的结果数据）；完全数据：观测数据X和隐随机变量的数据Z，对于分类模型，隐随机变量就是得到数据的类别。

3、设待估计的模型参数为θ，每个样本的真实是隐随机变量，未知；EM算法的步骤：

* 初始化
* E步：计算E(, E(可用表示。即计算在参数值为的情况下，样本真实类别的期望E(。

对于k-means，这一步计算的是当前聚类中心为，…,的条件下，样本的可能分类

* M步：E(代替带入L(θ)，求本轮迭代中使得极大似然函数最大的θ，即argL(θ)。对于k-means来说，即按照上一轮聚类中心将样本集划分后，将聚类中心更新，值为当前分类子集的质心。

1. EM推导：EM算法通过一步步迭代θ值，逐步最大化似然函数。假设在第n次迭代取值，更新θ值时，希望L(θ)-L()>0，以此逐步最小化似然函数，考虑两者的差，有Jensen不等式，当f(x)时凸函数时，此时f(x)=logx，满足不等式，且。
2. 算法的收敛性，观测数据的似然函数P(X|θ)在θ的迭代过程中单调递增。EM算法得到的θ值观测数据对数似然函数L(θ)=logP(X|θ) 的局部最优解。