《统计学习方法》第九章读书笔记

一、自己提出的问题的理解：

1. **提出的问题1：EM算法如何与无监督学习联系？**

**讨论后的理解**：EM算法可用于无监督学习模型。因为无监督学习没有标签（即y值），而EM算法中引入了隐变量的概念，可以先给无监督学习估计一个隐状态（即标签），观测数据当作样本值，算法模型就可以转换成有监督学习。

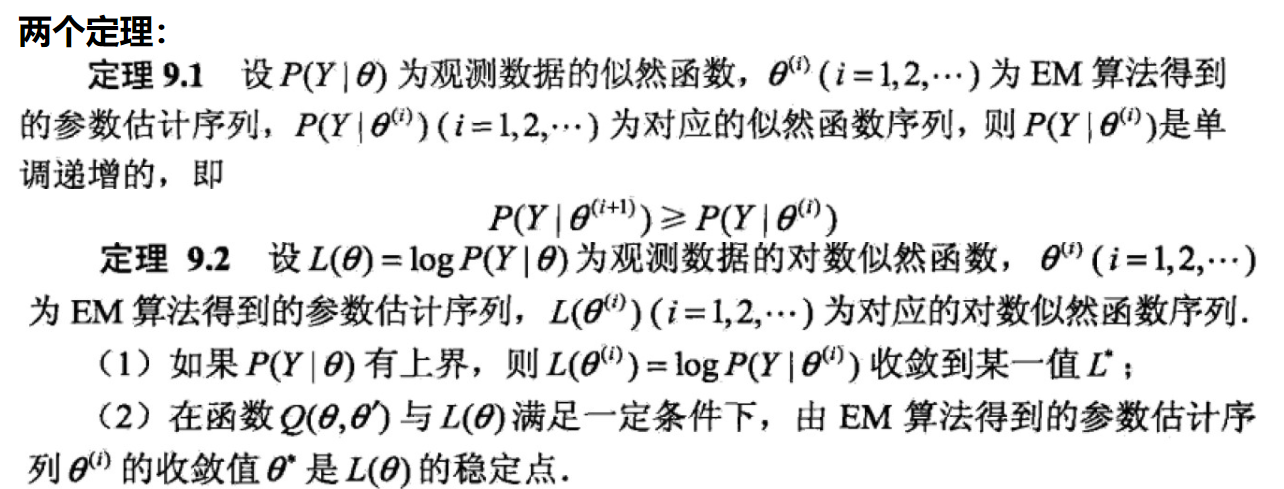
1. **提出的问题2：EM算法给出的局部最优解，这个“局部”是指哪个局部？什么情况下局部最优解可以是全局最优解？**

**讨论后的理解**：求出的是哪个局部最优解，与函数的形状和给定的初值有关。在样本规模不太大的情况下，可以对参数空间进行网格搜索，选取最优的那个解。当函数只有一个极大值点，是一个凸优化问题时，局部最优是全局最优。

二、别人提出的问题的理解：

1. 问题3：EM算法中，初值的选取重要吗？如何选取参数的初值？

**个人的理解：**EM算法的2个定理只能保证参数估计序列收敛到对数似然函数序列的稳定点，不能保证收敛到极大值点，所以初值选择十分重要。通常的做法是选取几个不同的初值进行迭代，然后对得到的各个估计值加以比较，从中选择最好的。



三、读书计划

1. 本周完成的内容章节：第九章
2. 下周计划：第十章

四、读书摘要及理解

 最大期望（EM）算法是在概率模型中寻找参数最大似然估计或者最大后验估计的算法，其中概率模型依赖于无法观测的隐性变量。最大期望算法经常用在机器学习和计算机视觉的数据聚类（Data Clustering）领域。最大期望算法经过两个步骤交替进行计算，第一步是计算期望（E），利用对隐藏变量的现有估计值，计算其最大似然估计值；第二步是最大化（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。

