一、自己的问题

1. EM算法中，初值的选取重要吗？如何选取参数的初值？

传统的EM算法对初始值敏感，聚类结果随不同的初始值而波动较大。总的来说，EM算法收敛的优劣很大程度上取决于其初始参数。通常的做法是选取几个不同的初值进行迭代，对得到的各个估计值加以比较，从中选择最好的。

2. EM算法有什么应用？

EM算法及其改进版本被用于机器学习算法的参数求解，常见的例子包括高斯混合模型、概率主成份分析、隐马尔可夫模型等非监督学习算法。EM算法可以给出隐变量，即缺失数据的后验 ，因此在缺失数据问题中有应用。

二、他人的问题

3、 EM算法如何与无监督学习联系？

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。

4、 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？

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。

三、读书计划

本周完成章节：第九章

下周计划：第十章

四、读书摘要

EM算法，是一类通过迭代进行[极大似然估计](https://baike.baidu.com/item/%E6%9E%81%E5%A4%A7%E4%BC%BC%E7%84%B6%E4%BC%B0%E8%AE%A1/3350286)的优化算法，通常作为[牛顿迭代法](https://baike.baidu.com/item/%E7%89%9B%E9%A1%BF%E8%BF%AD%E4%BB%A3%E6%B3%95/10887580)的替代用于对包含隐变量或缺失数据的[概率模型](https://baike.baidu.com/item/%E6%A6%82%E7%8E%87%E6%A8%A1%E5%9E%8B/9672385)进行参数估计。

EM算法的标准计算框架由E步和M步交替组成，算法的收敛性可以确保迭代至少逼近局部极大值。EM算法是MM算法的特例之一，有多个改进版本，包括使用了[贝叶斯推断](https://baike.baidu.com/item/%E8%B4%9D%E5%8F%B6%E6%96%AF%E6%8E%A8%E6%96%AD/833912)的EM算法、EM梯度算法、广义EM算法等。