**统计学习方法读报告**

09017244 郑健雄

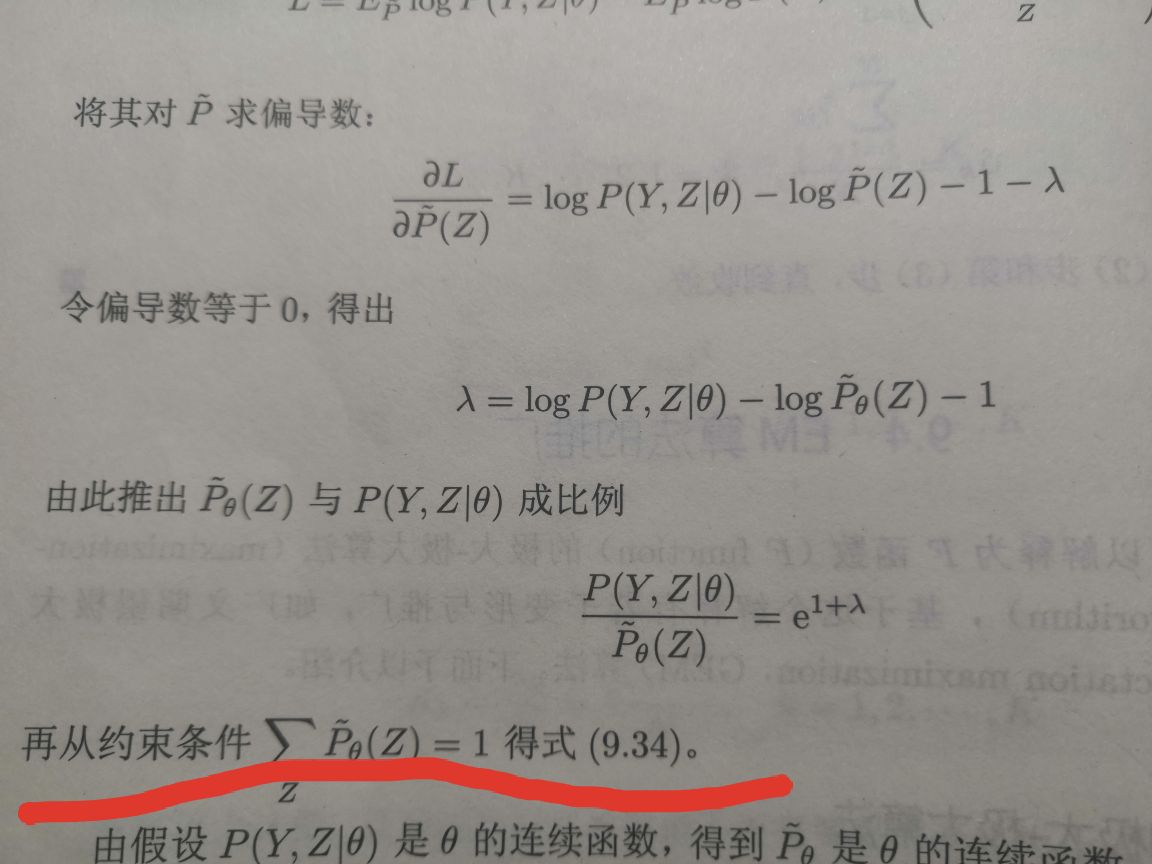
**本周进度:**统计学习方法第十章和第十一章内容，负责9.2讲解。

**下周进度:**统计学习方法第十二章和第十三章。

1. **问题与解答**

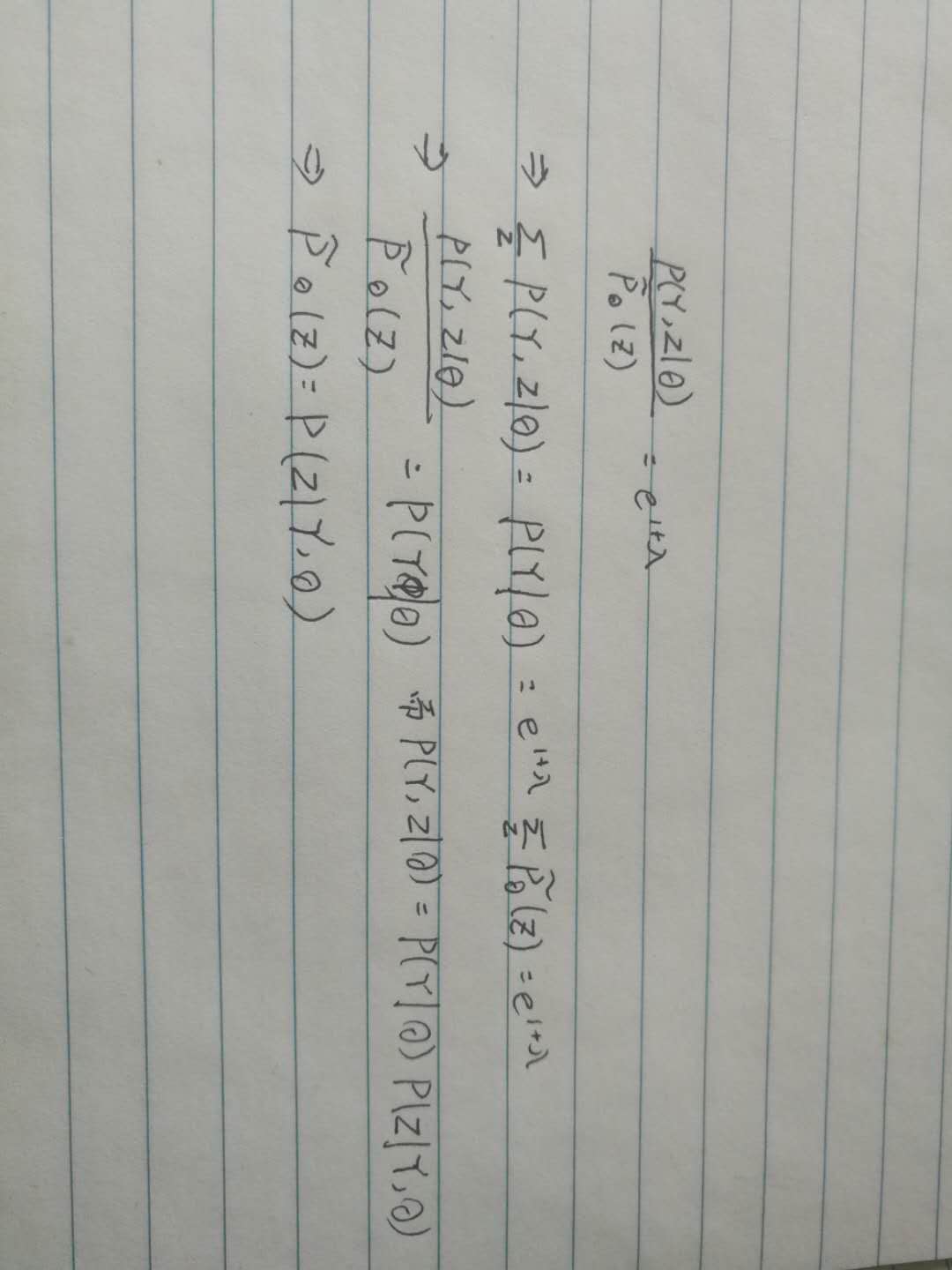
会上讨论的问题：

1.(别人提出)



这一步是如何推导的吗？

讨论结果:



经过以上推导以及概率论的知识可以得出结论。

2.(别人提出) EM算法的初值选择问题，不同初值结果不同，该如何选择初值？

讨论结果:

有很多方法可以提高EM算法的效果，一种比较简单的思路是随机选取一些初值，训练出来多个模型，选择一个泛化效果比较好的。

也可以使用讨论中说的，对各种初始值进行聚类分析，然后再聚类结束后选择聚类中心作为初始值，这样可以降低陷入局部极值的情况。

**二、 下周计划安排**

下周计划阅读完第十二章和第十三章。

**三、 读书收获**

第九章内容总结：

1. EM算法: EM算法是一种框架而非一种算法，其是为了处理含有隐变量的概率估计问题来提出的。其背后的原理本身是极大似然估计或者贝叶斯估计，但是因为有隐变量，所以需要同时进行隐变量的估计，其全名称为期望极大化算法，分为E(期望)和M(极大化)两步。

其中E步进行计算Q函数，也就是该算法的核心，而M步则极大化Q函数，得出模型参数。该模型的导出是通过极大似然函数的变形来得到的。

EM算法因为可以解决缺失数据问题，所以可以用于无标注数据生成标签问题和无监督学习问题。

2. EM算法具有收敛性:

根据证明，EM算法的Q函数时在迭代过程中是单调递增的，所以如果目标的概率分布有上界，而根据单调性，其一定会达到一个稳定点。

而在大多数条件下，Q函数的稳定点一般也是极大似然函数的稳定点。但是其一般只能保证到达一个局部最优点，而这个局部最优点不一定是全局最优的。

但是因为其具有收敛性，所以可以训练出一个较好的模型。

3.EM算法的实际应用:

比较典型的应用就是高斯混合模型的参数估计以及马尔可夫随机场。高斯混合模型的学习实际上就是学习多个线性相加的高斯分布的模型参数。其训练分为两步，E步的时候写出含参数的联合概率分布，而M步则通过最大化求出模型参数。

4. EM算法的推广:

EM算法实际上其本质可以理解为F函数的极大极大算法。在求解这个问题的过程中，其出现了Q函数的形式以及极大化的方法，而把这些求解方法中的结论提取出来，就产生了EM算法。

而GEM算法，实际上是一个更广泛的处理框架，其直接使用了F函数极大极大算法的思路作为基础而不是通用的EM算法，因而更具普遍性。一般有两种较为常用的GEM算法框架。通过逐步极大化Q函数的方法或者通过一次优化一个模型参数，进行多次条件极大化，得出新参数同时提升效率的算法。