方骏-2020年5月24日-读书报告

1. 自己提出的问题的理解：
2. 提出的问题1：定义松弛特征的目的是什么？

讨论后的理解： 因为在迭代尺度法中，要求表示数据中的特征总数，所以如果x不同，那么y不同，所以求出来的T值也不同，在计算中会带来麻烦，所以定义松弛特征，用一个比所有数据的特征和都大得多的常数来代替T，可以简化计算，而且由于是求最大值，所以具体T取什么值不影响。

1. 别人提出的问题的理解：
2. 读书计划

1、本周完成的内容章节：《统计机器学习》第十一章11.4-11.5

2、下周计划：《统计机器学习》第十二章

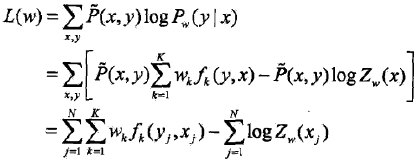
四、读书摘要及理解

1、条件随机场模型的学习方法包括极大似然估计和正则化的极大似然估计，优化实现算法有改进的迭代尺度发IIS、梯度下降法和拟牛顿法。

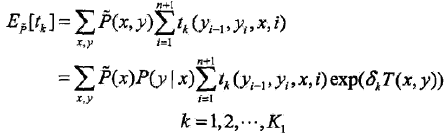
2、迭代尺度法要求极大化训练数据的对数似然函数来求模型参数：

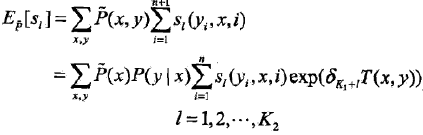


带入参数化的条件随机场模型公式，可得：



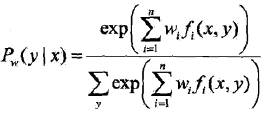
迭代尺度法就是不断优化对数似然函数改变量的下界，达到极大化对数似然函数的目的。求出参数向量的增量，通过增量更新参数向量。那么各特征的更新方程为：





在迭代的过程中，不断的求解上述过程的解，得到增量，然后更新权值，直到所有的权值都收敛。这里的T就是数据中的特征总数，见前面的问题。

3、拟牛顿法就是对于这种随机场模型：



学习的优化目标函数是：

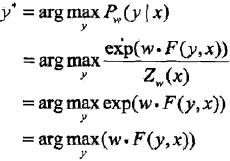


梯度函数是：

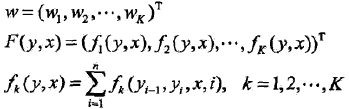


接着使用BFGS算法就可以了。

4、条件随机场的预测算法就是给定条件随机场和输入序列，求条件概率最大的输出序列，对观测序列进行标注。其实这个就是维特比算法，找到最优的路径：



这里的各个符号为：



写作：



最大值的路径写做：





接着套用维特比算法即可。