

# HOMEWORK 5

## 1. 证明题

(1) 证明《统计学习方法》习题 4.1:

通过极大似然方法估计先验概率和条件概率，注意写出极大似然法的目标函数及推导过程。

(2) 使用训练数据学习贝叶斯分类器，分别采用极大似然估计和贝叶斯估计 ( $\lambda = 1$ )，请同学们写出给定输入  $x = (2, M)^T$  时的类标记。注意写出推导过程。

|           | 1        | 2        | 3        | 4        | 5        | 6        | 7        | 8        | 9        | 10       | 11       | 12       | 13       | 14       | 15       |
|-----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| $X^{(1)}$ | 1        | 1        | 1        | 2        | 2        | 2        | 2        | 2        | 3        | 3        | 3        | 3        | 3        | 3        | 3        |
| $X^{(2)}$ | <i>S</i> | <i>M</i> | <i>L</i> | <i>S</i> | <i>M</i> | <i>M</i> | <i>L</i> | <i>S</i> | <i>L</i> | <i>L</i> | <i>S</i> | <i>M</i> | <i>M</i> | <i>M</i> | <i>L</i> |
| $Y$       | 1        | 1        | -1       | 1        | -1       | -1       | -1       | 1        | 1        | 1        | -1       | 1        | 1        | 1        | 1        |

## 2. 市长电话数据分析及算法实现

编程语言可以使用 R/python，推荐使用 R 语言，提交 rmarkdown 输出的报告。具体任务见 word 文档。

最后以 HTML/PDF 的形式提交报告。报告中需包括题目内容中涉及的代码和相关文字解释、结果分析。

提交时间：11 月 30 日，晚 20:00 之前。请预留一定的时间，迟交作业扣 3 分，作业抄袭 0 分。