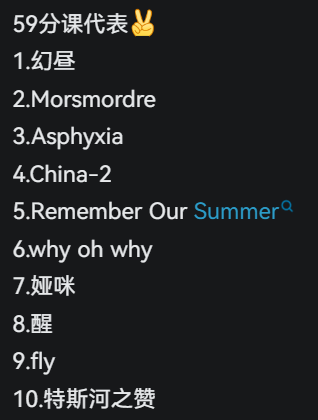
第14次作业

1.

(1) 介绍一下 EM算法



EM算法是一种迭代优化算法，用于求解包含隐变量的概率模型参数的最大似然估计或最大后验概率估计。EM算法的全称是Expectation-Maximization Algorithm（期望最大化算法）。

通常用于解决存在隐变量（或者称为缺失数据）的统计模型参数估计问题。在这种情况下，由于存在隐变量，无法直接应用传统的最大似然估计方法。EM算法通过迭代的方式，通过两个步骤交替进行：E步骤（Expectation Step，期望步骤）和M步骤（Maximization Step，最大化步骤），交替进行，直到收敛到某个阈值或达到最大迭代次数为止。

(2) 设计不同类用户A，B，数量不均，其消费日均数分别在300和1000，通过EM算法估算两用户的平均日均消费

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Col a | Col b | Col c | Col d | Col e | Col f |
| Idx 1 |  |  |  |  |  |  |
| Idx 2 |  | 77 |  |  | 89 |  |
| Idx 3 |  |  | 56 |  |  |  |
| Idx 4 | 23 |  |  |  |  | 8 |
| Idx 5 |  |  |  |  | 23 |  |

图表 1