

EM 算法实验内容

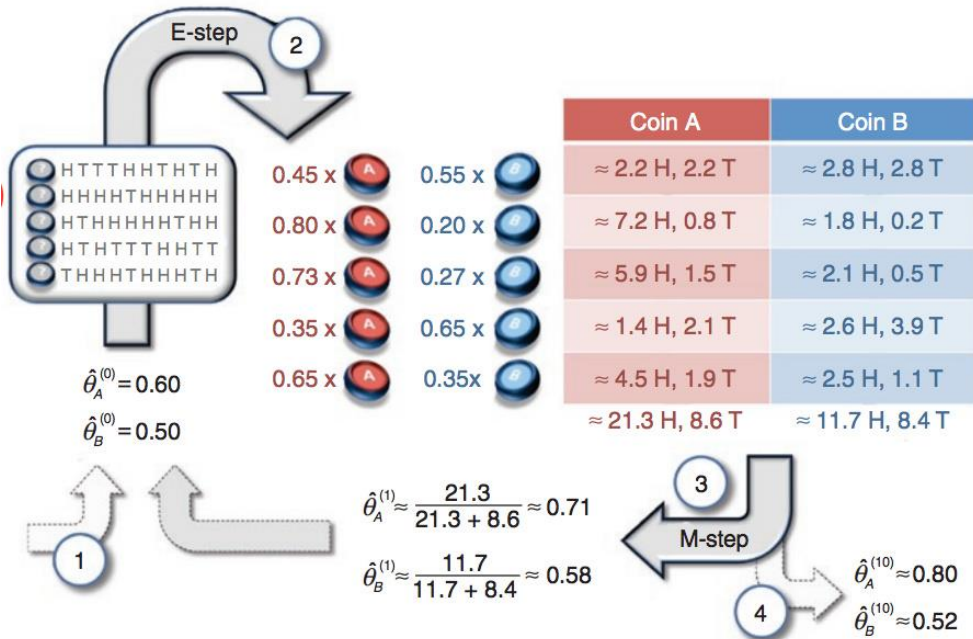
假设有两枚硬币 A、B，以相同的概率随机选择一个硬币，进行如下的抛硬币实验：共做 5 次实验，每次实验独立的抛十次，结果如图中 a 所示，例如某次实验产生了 H、T、T、T、H、H、T、H、T、H，H 代表正面朝上。

假设试验数据记录员可能是实习生，业务不一定熟悉，造成如下图的 a 和 b 两种情况：

a Maximum likelihood



b Expectation maximization



a 表示实习生记录了详细的试验数据，我们可以观测到试验数据中每次选择的是 A 还是 B

b 表示实习生忘了记录每次试验选择的是 A 还是 B，我们无法观测实验数据中选择的硬币是哪一个

问：在两种情况下分别如何估计两个硬币正面出现的概率？并完成相应的 python 代码实现。

图片分类任务

按要求对 CIFAR-10 数据集使用多种分类器进行图片分类任务。作业要求：

- 明确图片分类任务的基本流程
- 清楚数据集的训练集验证集和测试集的划分和用途，简单介绍自己的用法
- 编写 K 近邻算法，SVM 以及简单的两层神经网络分类器各一（多分类分类器）进行图片分类任务
- 根据得到的分类结果（精度）说明和比较各个算法的优越性和局限性，懂得它们之间的差异
- 编写实验报告

数据集下载链接

<http://www.cs.toronto.edu/~kriz/cifar.html>