

**Problem 1**

**3.7** 令码长为 9，类别为 4，试给出海明距离意义下理论最优的 ECOC 二元码并证明之。

**Solution:**

数量的咖啡机数量的会计法

**Problem 2**

**3.8** ECOC 编码能起到理想纠错作用的重要条件是：在每一位编码上出错的概率相当且独立。试析多分类任务经 ECOC 编码后产生的二类分类器满足该条件的可能性及由此产生的影响。

**Problem 3**

**3.9** 使用 OvR 和 MvM 将多分类任务分解为二分类任务求解是，试述为何无需专门针对类别不平衡性进行处理。

**Solution:**

对于 OvR 和 MvM 而言，每次的数据集都是不均衡的，但是其会遍历每一个类，这样类别的不均衡性在遍历的过程中就抵消了。比如一共有 4 类， $c_1, c_2, c_3, c_4$ ，第一次取  $c_1$  为正样本，其余为负样本，这时正样本  $c_1$  的数目就少，但是第二次将  $c_2$  作为正样本时， $c_1$  就变成了负样本，这样对于分类器  $c_1$  的样本又变多了，所以不断地遍历，使得每一种样本都能够被取到，这样就不用进行不平衡处理。

**Problem 4**

**3.10** 试推导多分类代价敏感学习（仅考虑基于类别的误分类代价）使用“在放缩”能获得理论最优解的条件。