14.4 比较k均值聚类与高斯混合模型加EM算法的异同。

相似之处：

①目标：K均值聚类和高斯混合模型加EM算法都用于无监督学习，将数据集中的样本划分为不同的聚类。

②聚类数量：两种方法都需要预先指定聚类的数量。

不同之处：

①假设：K均值聚类假设每个聚类的样本点属于同一个簇，即样本点在特征空间中的分布是球状的，并且每个聚类的中心点是特征空间中样本点的平均值。而高斯混合模型假设每个聚类的样本点服从高斯分布，即样本点在特征空间中的分布是椭圆形的，并且每个聚类的中心点是高斯分布的均值。

②数据点分配：在K均值聚类中，每个数据点只能被分配给一个聚类，而在高斯混合模型中，每个数据点可以被分配给多个聚类，并计算其属于每个聚类的概率。

③簇的形状：由于K均值聚类假设样本点的分布是球状的，因此对于非球状的聚类，K均值聚类可能表现不佳。而高斯混合模型可以处理各种形状的聚类，因为它使用高斯分布建模。

④对初始参数的敏感性：K均值聚类对初始的聚类中心点选择非常敏感，不同的初始选择可能导致不同的结果。而高斯混合模型加EM算法对初始参数的选择不敏感，通常可以使用随机初始化。

⑤概率表示：高斯混合模型给出了数据点属于每个聚类的概率，可以用于后续的概率推断和决策，而K均值聚类只提供了数据点的硬聚类结果。