计算过程：
已知：输入数据集 d 中有 9 个元素。d1, d2, …, d9。
问题：对数据集进行聚类。

K-means 聚类方法：
设 c1, c2, c3, c4 为用于循环的变量。(c1, c2, c3, c4)的序号对应着簇的编号。
变量命名规则：
c 表示簇中心，center。
ct 表示 c test，用于误差的判断过程中。

1. 随便挑选数据集 d 中的 4 个元素(c1, c2, c3, c4)给赋初始值。
2. 计算数据集中的所有元素分别到 c1, c2, c3, c4 的距离，并将记录和用于计算的元素距离最小的簇的编号。（这一步完成，d 中所有元素都会有一个簇号。）
3. 依次对各簇中的元素取平均，得到判断变量 ct1, ct2, ct3, ct4。
4. 计算 c 和 ct 的误差。就是根据(c1, c2, c3, c4) 和(ct1, ct2, ct3, ct4)，依次计算 c1-ct1, c2-ct2, c3-ct3, c4-ct4。
5. 判断误差是否达到收敛要求。如果达到收敛要求，分类完成；如果没有达到，将(ct1, ct2, ct3, ct4) 赋值给 (c1, c2, c3, c4)，进入第 2 步。（ct 向 c 的赋值其实就是第 1 步。）

循环变量的个数对应着簇的个数，也即类的个数。可以设置不同的循环变量个数，循环变量的取值等于簇的数量，它的大小就是“K-means”中 K 的大小。
K-means 是基于空间距离的分类方法。K-means 需要预先设定簇的数量，即 K 值。有的时候，在分类之前，我们并不清楚数据将被分为多少类别，只能不断试验 K 值。