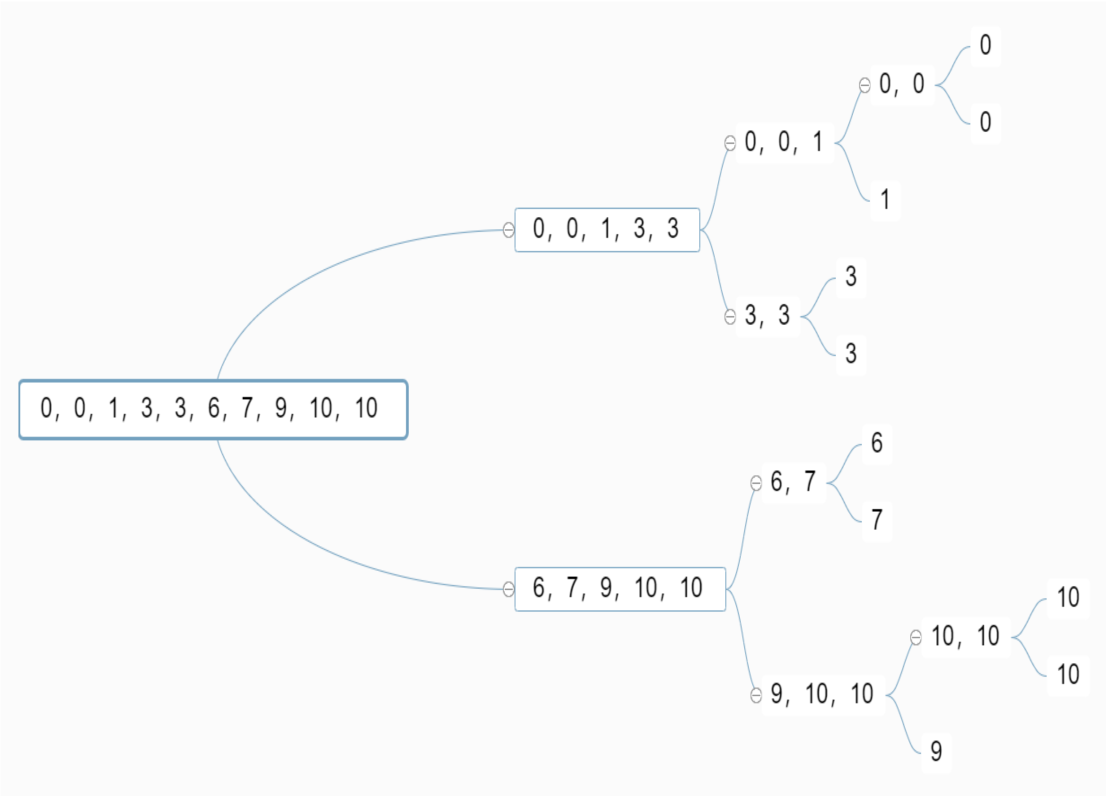
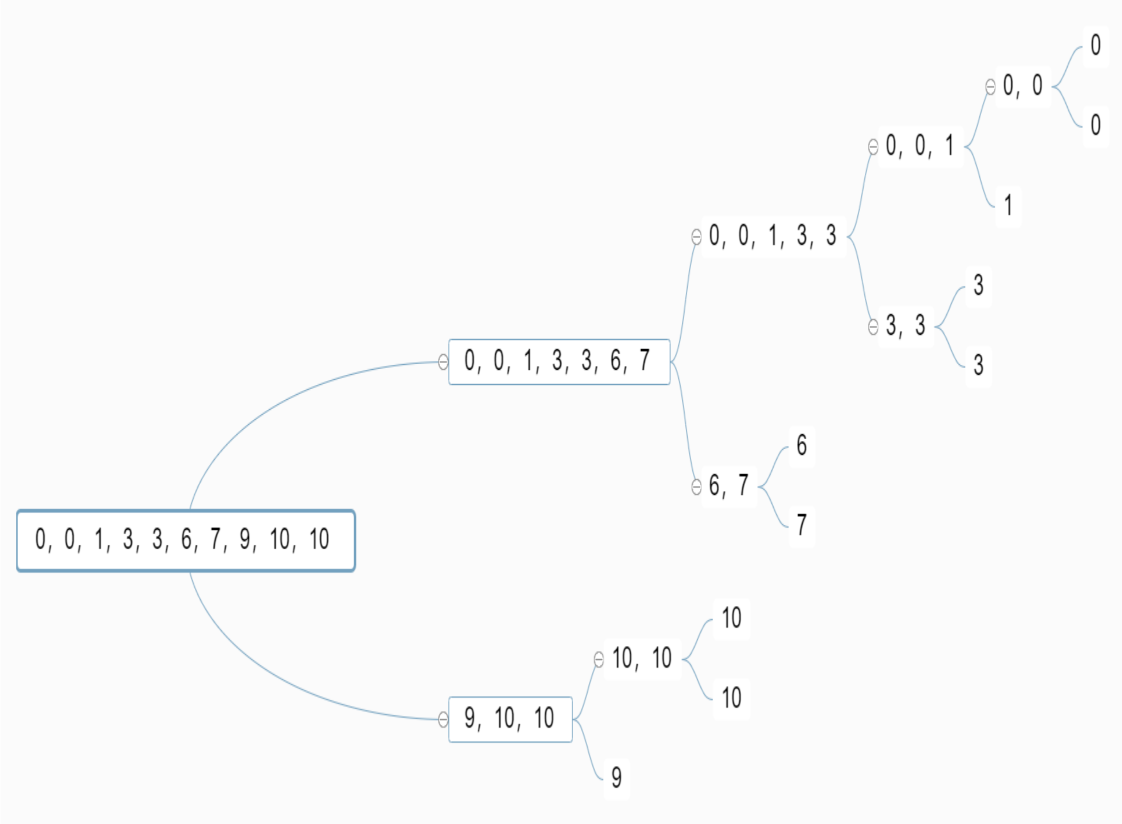
第 19 章

3：



0，0，1，3，3，6，7，9，10，10
0，0，1，3，3
0，0，1
0，0
0
0
1
3，3
3
3
6，7，9，10，10
6，7
6
7
9，10，10
10，10
10
10
9

4：



7：

算法每次执行时，都依据最小欧式距离的准则进行簇划分，并寻找新的簇质心，通过不断更新簇质心的位置与分簇，在这个过程中，使得簇内差异不断减小，簇间差异不断增大。

给定 n 个训练样本{x1，x2，x3，...，xn}

算法过程描述如下所示：

1.创建 k 个点作为起始质心点，c1，c2，...，ck

2.重复以下过程直到收敛

遍历所有样本 xi

遍历所有质心 cj

记录质心与样本间的距离

将样本分配到距离其最近的质心

对每一个类，计算所有样本的均值并将其作为新的质心

闵可夫斯基距离(Minkowski Distance)：计算距离的通用的公式：

$$d(i,j)=\sqrt[h]{|x_{i1}-x_{j1}|^h+|x_{i2}-x_{j2}|^h+\ldots+|x_{ip}-x_{jp}|^h}$$

i = (xi1, xi2, …, xip) 和 j = (xj1, xj2, …, xjp) 是 p 维数据对象

曼哈顿距离(或城市块距离 Manhattan distance):h=1

$$d(i,j)=|x_{i1}-x_{j1}|+|x_{i2}-x_{j2}|+\ldots+|x_{ip}-x_{jp}|$$

欧几里德距离(用的最多的):h=2

$$d(i,j)=\sqrt{|x_{i1}-x_{j1}|^2+|x_{i2}-x_{j2}|^2+\ldots+|x_{ip}-x_{jp}|^2}$$

11：

第 2 遍：

$$\mathrm{SSB} = \sum_{i}^{k} n_i \bullet d\left(m_i, M\right)^2$$

$$= 4 \bullet d\left(\left(1.25, 1.75\right), \left(2.6525, 2.25\right)\right)^2 + 4 \bullet d\left(\left(4, 2.75\right), \left(2.6525, 2.25\right)\right)^2$$

$$= 17.125$$

$$\mathrm{MSB} = \frac{\mathrm{SSB}}{\mathrm{k} - 1} = 17.131$$

$$\mathrm{SSE} = \sum_{i}^{k} \sum_{\mathrm{p} \in \mathrm{C}_i} d\left(p, m_i\right)^2$$

$$= 1^2 + 0.85^2 + 0.72^2 + 1.52^2 + 0^2 + 0.57^2 + 1^2 + 1.41^2$$

$$= 7.8643$$

$$\mathrm{MSE} = \frac{\mathrm{SSE}}{\mathrm{N} - \mathrm{k}} = \frac{7.8643}{6} = 1.3107$$

$$\mathrm{F} = \frac{\mathrm{MSB}}{\mathrm{MSE}} = 13.07$$

第 3 遍

$$\mathrm{SSB} = \sum_{i}^{k} n_i \bullet d\left(m_i, M\right)^2$$

$$= 4 \bullet d\left(\left(1.25, 1.75\right), \left(2.6525, 2.25\right)\right)^2 + 4 \bullet d\left(\left(4, 2.75\right), \left(2.6525, 2.25\right)\right)^2$$

$$= 17.131$$

$$\mathrm{MSB} = \frac{\mathrm{SSB}}{\mathrm{k} - 1} = 17.131$$

$$\mathrm{SSE} = \sum_{i}^{k} \sum_{\mathrm{p} \in \mathrm{C}_i} d\left(p, m_i\right)^2$$

$$= 1.27^2 + 1.03^2 + 0.25^2 + 1.03^2 + 0.35^2 + 0.75^2 + 0.79^2 + 1.06^2$$

$$= 6.2299$$

$$\mathrm{MSE} = \frac{\mathrm{SSE}}{\mathrm{N} - \mathrm{k}} = \frac{6.2299}{6} = 1.0383$$

$$\mathrm{F} = \frac{\mathrm{MSB}}{\mathrm{MSE}} = 16.499$$