

## 6.1

初始两个质心为(0, 4), (3, 3)

第一次迭代

(1, 3), (1, 4)属于第一簇, (2, 3), (2, 4), (3, 1), (3, 2), (4, 1), (4, 2)属于第二簇

质心更新为 $(1, \frac{7}{2})$ ,  $(3, \frac{13}{6})$

第二轮迭代

(1, 3), (1, 4), (2, 3), (2, 4)属于第一簇

(3, 1), (3, 2), (4, 1), (4, 2)属于第二簇

质心更新为 $(\frac{3}{2}, \frac{7}{2})$ ,  $(\frac{7}{2}, \frac{3}{2})$

第三轮迭代

没有簇样本所属类发生了改变, 质心不变, 停止迭代

所以, 最终的聚类中心为 $(\frac{3}{2}, \frac{7}{2})$ ,  $(\frac{7}{2}, \frac{3}{2})$

## 6.2

如果样本集合为凸集则一定能得到最终不会改变的质心, 否则的话可以通过设置迭代次数或者设定收敛判断距离来停止迭代

## 6.3

$$X = \begin{bmatrix} 2 & 3 & 3 & 4 & 5 & 7 \\ 2 & 4 & 5 & 5 & 6 & 8 \end{bmatrix}$$

根据协方差矩阵的公式

$$S = [s_{ij}]_{m \times m}$$
$$s_{ij} = \frac{1}{n-1} \sum_{k=1}^n (x_{ik} - \bar{x}_i)(x_{jk} - \bar{x}_j)$$

所以X的相关矩阵为

$$R = \begin{bmatrix} 1 & 0.95032889 \\ 0.95032889 & 1 \end{bmatrix}$$

对R进行特征值分解, 得到相关矩阵的特征值, 和对应的特征向量

$$\lambda_1 = 0.04967111, \alpha_1 = [0.70710678 \quad -0.70710678]$$

$$\lambda_2 = 1.95032889, \alpha_2 = [0.70710678 \quad +0.70710678]$$

所以

$$\begin{aligned}y_1 &= 0.70710678x_1 - 0.70710678x_2 \\y_2 &= 0.70710678x_1 + 0.70710678x_2\end{aligned}$$

## 6.4

A,B

####

## 6.5

D

## 6.6

D

## 6,7

D

## 6.8

初始随机选择两个点作为质心为 $(0, 0)$ ,  $(5, 0)$

第一次迭代

$(0, 0)$ ,  $(1, 0)$ ,  $(0, 2)$ 属于第一簇,  $(5, 0)$ ,  $(5, 2)$ 属于第二簇

质心更新为 $(\frac{1}{3}, \frac{2}{3})$ ,  $(5, 1)$

第二轮迭代

没有簇样本所属类发生了改变, 质心不变, 停止迭代

所以, 最终的聚类中心为 $(\frac{1}{3}, \frac{2}{3})$ ,  $(5, 1)$

$(0, 0)$ ,  $(1, 0)$ ,  $(0, 2)$ 属于第一簇,  $(5, 0)$ ,  $(5, 2)$ 属于第二簇