

1.什么叫做最小距离分类器？描述几种最小距离分类器的特点，并说明其适用情况，可画图示意

解：

最小距离分类器：利用两个模式在特征空间中的距离作为两者之间的相似性度量，由此设计的度量函数就是最小距离分类器。

基于标准样本的距离分类器、基于平均样本法的距离分类器、基于平均距离法的距离分类器、基于最近邻法的距离分类器

2.对于某样本集合：

1) 何为总体线性可分？何为成对线性可分？

2) 在成对线性可分情况下，如何避免最终得到的决策域中不存在不确定区域？简要说明其原理；

3) 如果样本集成对线性可分，是否一定总体线性可分？反之如何？可画图说明。

解：

1) 总体线性可分：若能用一个超平面将任意一个类别的样本同其他类别的样本分开，则称该样本集合是总体线性可分的
成对线性可分：若能用一个超平面将所有类别中任意两个类别的样本分开，则称该样本集合为成对线性可分的。

2) 为了保证最终得到的决策域中不存在不确定区域，采用如下策略：对于每个类别定义一个判别函数，并采用最大值判决准则完成分类。

3) 若样本几何总体线性可分，则也一定成对线性可分，反之未必。

3.什么是 k 近邻法？给定两类训练样本如下：

ω_1 : (1,0), (3,0), (3,1), (2,-1)

ω_2 : (-1,1), (2,0), (1,-1), (4,1), (2,2)

请采用 k 近邻法（分别取 k=3 和 5），计算测试样本(1,1)的分类结果。

解：

k=3: ω_1 : (1,0) ω_2 : (2,0) (2,2)

(1,1)分类为 ω_2

k=5: ω_1 : (1,0) (3,1) ω_2 : (2,0) (2,2) (-1,1)

(1,1)分类为 ω_2

4.位势函数法相较 H-K 算法的优点是? 利用位势法对以下模式类进行分类, 计算位势函数。

$\omega_1: (0,0)^T, (1,2)^T$ $\omega_2: (1,-1)^T, (3,0)^T$ (位势函数选用 $K(X, X_n) = e^{-\|X-X_n\|^2}, K_{A,0}(X)=0$)

解:

(1)优点是位势法可以用于非线性可分的情况。

(2)计算步骤:

$$k=0, X'=X_1=(0,0)^T, K_{A,0}(X')=0, N_c=0$$

$$\text{因为 } X' \in \omega_1, K_{A,0}(X') \leq 0$$

$$K_{A,1}(X) = K_{A,0}(X) + K(X, X') = e^{-\|X\|^2}$$

$$k=1, X'=X_2=(1,2)^T, K_{A,1}(X') = e^{-5} > 0, N_c=1$$

$$\text{因为 } X' \in \omega_1, K_{A,1}(X') > 0$$

$$K_{A,2}(X) = K_{A,1}(X) = e^{-\|X\|^2}$$

$$k=2, X'=X_3=(1,-1)^T, K_{A,2}(X') = e^{-2} > 0, N_c=0$$

$$\text{因为 } X' \in \omega_2, K_{A,2}(X') > 0$$

$$K_{A,3}(X) = K_{A,2}(X) - K(X, X') = e^{-\|X\|^2} - e^{-\|X-(1,-1)^T\|^2}$$

$$k=3, X'=X_4=(3,0)^T, K_{A,3}(X') = e^{-9} - e^{-5} < 0, N_c=1$$

$$\text{因为 } X' \in \omega_2, K_{A,3}(X') < 0$$

$$K_{A,4}(X) = K_{A,3}(X) = e^{-\|X\|^2} - e^{-\|X-(1,-1)^T\|^2}$$

$$k=4, X'=X_1=(0,0)^T, K_{A,4}(X') = 1 - e^{-2} > 0, N_c=2$$

$$\text{因为 } X' \in \omega_1, K_{A,4}(X') > 0$$

$$K_{A,5}(X) = K_{A,4}(X) = e^{-\|X\|^2} - e^{-\|X-(1,-1)^T\|^2}$$

$$k=5, X'=X_2=(1,2)^T, K_{A,5}(X') = e^{-5} - e^{-9} > 0, N_c=3$$

$$\text{因为 } X' \in \omega_1, K_{A,5}(X') > 0$$

$$K_{A,6}(X) = K_{A,5}(X) = e^{-\|X\|^2} - e^{-\|X-(1,-1)^T\|^2}$$

$$k=6, X'=X_3=(1,-1)^T, K_{A,6}(X') = e^{-2} - 1 < 0, N_c=4$$

$$\text{因为 } N_c=4, \text{ 迭代结束, 最终位势函数为 } e^{-\|X\|^2} - e^{-\|X-(1,-1)^T\|^2}。$$

5、有三个模式类： $\omega_1: X_1 = (-1, -1)^T$, $\omega_2: X_2 = (2, -1)^T$, $\omega_3: X_3 = (-2, 2)^T$

(1) 是否总体线性可分？是否成对线性可分？

(2) 采用感知器算法设计合适的分类器 ($W(0) = \mathbf{0}$, $\rho = 1$)，请给出相应的迭代过程。

答：

(1) 既是总体线性可分，又是成对线性可分。

(2) 见教材 P31-P33，使用三种情况的解答都可以，但推荐采用第三种情况，可以得到没有不确定区域的解。注意，第三种情况采用的是最大值判决。

模式类的增广形式： $\omega_1: X_1 = (-1, -1, 1)^T$, $\omega_2: X_2 = (2, -1, 1)^T$, $\omega_3: X_3 = (-2, 2, 1)^T$

三个决策函数： $g_1(X) = W_1^T X$, $g_2(X) = W_2^T X$, $g_3(X) = W_3^T X$

初始权向量： $W_1(0) = W_2(0) = W_3(0) = [0, 0, 0]^T$, $\rho = 1$

1) $k = 0$, $g_1(X_1) = 0$, $g_2(X_1) = 0$, $g_3(X_1) = 0$, $N_c = 0$

由于 $g_1(X_1) = g_2(X_1)$, $g_1(X_1) = g_3(X_1)$, 所以对 $W_1 W_2 W_3$ 进行修正

$$W_1(1) = W_1(0) + \rho X_1 = (-1, -1, 1)^T$$

$$W_2(1) = W_2(0) - \rho X_1 = (1, 1, -1)^T$$

$$W_3(1) = W_3(0) - \rho X_1 = (1, 1, -1)^T$$

2) $k = 1$, $g_1(X_2) = 0$, $g_2(X_2) = 0$, $g_3(X_2) = 0$, $N_c = 0$

由于 $g_2(X_2) = g_1(X_2)$, $g_2(X_2) = g_3(X_2)$, 所以对 $W_1 W_2 W_3$ 进行修正

$$W_1(2) = W_1(1) - \rho X_2 = (-3, 0, 0)^T$$

$$W_2(2) = W_2(1) + \rho X_2 = (3, 0, 0)^T$$

$$W_3(2) = W_3(1) - \rho X_2 = (-1, 2, -2)^T$$

3) $k = 2$, $g_1(X_3) = 6$, $g_2(X_3) = -6$, $g_3(X_3) = 4$, $N_c = 0$

由于 $g_1(X_3) > g_3(X_3)$, 所以对 $W_1 W_3$ 进行修正

$$W_1(3) = W_1(2) - \rho X_3 = (-1, -2, -1)^T$$

$$W_2(3) = W_2(2) = (3, 0, 0)^T$$

$$W_3(3) = W_3(2) + \rho X_3 = (-3, 4, -1)^T$$

4) $k = 3$, $g_1(X_1) = 2$, $g_2(X_1) = -3$, $g_3(X_1) = -2$, $N_c = 1$

由于 $g_1(X_1) > g_2(X_1)$, $g_1(X_1) > g_3(X_1)$, 所以无需修正

$$W_1(4) = W_1(3) = (-1, -2, -1)^T$$

$$W_2(4) = W_2(3) = (3, 0, 0)^T$$

$$W_3(4) = W_3(3) = (-3, 4, -1)^T$$

5) $k = 4$, $g_1(X_2) = -1$, $g_2(X_2) = 6$, $g_3(X_2) = -11$, $N_c = 2$

由于 $g_2(X_2) > g_1(X_2)$, $g_2(X_2) > g_3(X_2)$, 所以无需修正

$$W_1(5) = W_1(4) = (-1, -2, -1)^T$$

$$W_2(5) = W_2(4) = (3, 0, 0)^T$$

$$W_3(5) = W_3(4) = (-3, 4, -1)^T$$

6) $k = 5$, $g_1(X_3) = -3$, $g_2(X_3) = -6$, $g_3(X_3) = 13$, $N_c = 3$

由于 $g_3(X_3) > g_1(X_3)$, $g_3(X_3) > g_2(X_3)$, 所以无需修正

$$W_1(6) = W_1(5) = (-1, -2, -1)^T$$

$$W_2(6) = W_2(5) = (3, 0, 0)^T$$

$$W_3(6) = W_3(5) = (-3, 4, -1)^T$$

由于 $N_c = 3$, 迭代终止

得到判决函数

$$g_1(X) = W_1^T X = [-1, -2, -1]X$$

$$g_2(X) = W_2^T X = [3, 0, 0]X$$

$$g_3(X) = W_3^T X = [-3, 4, -1]X$$

决策面函数：

$$G_{12}(X) = g_1(X) - g_2(X) = [-4, -2, -1]X$$

$$G_{13}(X) = g_1(X) - g_3(X) = [2, -6, 0]X$$

$$G_{23}(X) = g_2(X) - g_3(X) = [6, -4, 1]X$$

6.若准则函数的形式为 $J(W, X) = \frac{1}{2} (1 - W^T X)^2$, $W = (1, 1)^T$, $X_1 = (1, 2)^T$, $X_2 = (-2, 1)^T$ 。

在应用梯度下降法时, 求 $\nabla J(W, X_1)$, $\nabla J(W, X_2)$ 。

解:

$$\nabla J(W, X) = \frac{\partial J}{\partial W} = (1 - W^T X)(-X)$$

$$\nabla J(W, X_1) = \frac{\partial J}{\partial W} = (1 - W^T X_1)(-X_1) = (-2)(-1, -2)^T = (2, 4)^T$$

$$\nabla J(W, X_2) = \frac{\partial J}{\partial W} = (1 - W^T X_2)(-X_2) = (2)(2, -1)^T = (4, -2)^T$$

7.有两个模式类: $\omega_1: (0,0)^T, (1,2)^T$; $\omega_2: (1,-1)^T, (3,0)^T$, 试用 Fisher 线性判别法构造分类器, 确定最佳投影方向, 并简要说明类内总离散度矩阵的含义

解:

(1) 计算两个样本的均值向量:

$$m_x^1 = \frac{1}{n_1} \sum_{X_k \in \omega_1} X_k = [0.5, 1]^T$$

$$m_x^2 = \frac{1}{n_2} \sum_{X_k \in \omega_2} X_k = [2, -0.5]^T$$

(2) 计算类内总离散度矩阵:

$$\begin{aligned} S_w &= \sum_{X_k \in \omega_1} (X_k - m_x^1)(X_k - m_x^1)^T + \sum_{X_k \in \omega_2} (X_k - m_x^2)(X_k - m_x^2)^T \\ &= \begin{bmatrix} 0.5 & 1 \\ 1 & 2 \end{bmatrix} + \begin{bmatrix} 2 & 1 \\ 1 & 0.5 \end{bmatrix} = \begin{bmatrix} 2.5 & 2 \\ 2 & 2.5 \end{bmatrix} \end{aligned}$$

(3) 计算 S_w 的逆矩阵 S_w^{-1} :

$$S_w^{-1} = \begin{bmatrix} \frac{10}{9} & -\frac{8}{9} \\ -\frac{8}{9} & \frac{10}{9} \end{bmatrix}$$

(4) 计算最佳投影

$$W^* = S_w^{-1}(m_x^1 - m_x^2) = [-3, 3]^T$$

(5) 类内总离散度矩阵刻画了投影后两类样本偏离各自均值的总的差异程度。

1. 说明最大后验概率判决准则为什么可以被成为最小错误概率判决准则。

最大后验概率判决:

$$p(X|\omega_1)P(\omega_1) > p(X|\omega_2)P(\omega_2) \Rightarrow X \in \omega_1$$

$$p(X|\omega_1)P(\omega_1) < p(X|\omega_2)P(\omega_2) \Rightarrow X \in \omega_2$$

因此, 决策面为

$$p(X|\omega_1)P(\omega_1) = p(X|\omega_2)P(\omega_2)$$

在课本图 3.2 中对应于图(c)的情形。

而错误概率:

$$P(e) = P(x \text{ 在 } \Omega_2 \text{ 中} | \omega_1) + P(x \text{ 在 } \Omega_1 \text{ 中} | \omega_2) = \int_{\Omega_2} p(x|\omega_1)P(\omega_1)dx + \int_{\Omega_1} p(x|\omega_2)P(\omega_2)dx$$

2. 已知两个一维模式类别的类概率密度函数为

$$p(x|\omega_1) = \begin{cases} x, & 0 \leq x \leq 1 \\ 2-x, & 1 \leq x \leq 2 \\ 0, & \text{else} \end{cases}$$

$$p(x|\omega_2) = \begin{cases} x-1, & 1 \leq x < 2 \\ 3-x, & 2 \leq x \leq 3 \\ 0, & \text{else} \end{cases}$$

先验概率分别为 $p(\omega_1)=0.4$, $p(\omega_2)=0.6$ 。试求最大后验概率判决函数以及总的分类错误概率 $P(e)$ 。

$$p(X|\omega_1)P(\omega_1) > p(X|\omega_2)P(\omega_2) \Rightarrow X \in \omega_1$$

$$p(X|\omega_1)P(\omega_1) < p(X|\omega_2)P(\omega_2) \Rightarrow X \in \omega_2$$

$$p(x|\omega_1)P(\omega_1) = \begin{cases} 0.4x, & 0 \leq x < 1 \\ 0.8-0.4x, & 1 \leq x \leq 2 \\ 0, & \text{else} \end{cases}$$

$$p(x|\omega_2)P(\omega_2) = \begin{cases} 0.6x-0.6, & 1 \leq x < 2 \\ 1.8-0.6x, & 2 \leq x \leq 3 \\ 0, & \text{else} \end{cases}$$

$0 < x < 1$:

$$p(x|\omega_1)P(\omega_1) = 0.4x$$

$$p(x|\omega_2)P(\omega_2) = 0$$

$$p(x|\omega_1)P(\omega_1) > p(x|\omega_2)P(\omega_2) \Rightarrow x \in \omega_1$$

$1 \leq x < 2$:

$$p(x|\omega_1)P(\omega_1) = 0.8-0.4x$$

$$p(x|\omega_2)P(\omega_2) = 0.6x-0.6$$

$$s.t. p(x|\omega_1)P(\omega_1) = p(x|\omega_2)P(\omega_2) \Rightarrow x = 1.4$$

$1 \leq x < 1.4$:

$$p(x|\omega_1)P(\omega_1) > p(x|\omega_2)P(\omega_2) \Rightarrow x \in \omega_1$$

$1.4 < x < 2$:

$$p(x|\omega_1)P(\omega_1) < p(x|\omega_2)P(\omega_2) \Rightarrow x \in \omega_2$$

$2 \leq x < 3$:

$$p(x|\omega_1)P(\omega_1) = 0$$

$$p(x|\omega_2)P(\omega_2) = 1.8 - 0.6x$$

$$p(x|\omega_1)P(\omega_1) < p(x|\omega_2)P(\omega_2) \Rightarrow x \in \omega_2$$

综上,

$$\begin{cases} \omega_1, & 0 < x < 1.4 \\ \omega_2, & 1.4 < x < 3 \\ \text{无法判断,} & \text{otherwise} \end{cases}$$

分类错误概率:

$$\begin{aligned} P(e) &= \int_0^{1.4} p(x|\omega_2)P(\omega_2)dx + \int_{1.4}^3 p(x|\omega_1)P(\omega_1)dx \\ &= \int_1^{1.4} (0.6x - 0.6)dx + \int_{1.4}^2 (0.8 - 0.4x)dx = 0.12 \end{aligned}$$

3. 从似然比形式的角度说明最大后验概率判决准则同最小风险判决准则之间的联系与区别。

注意: 似然比为相应两个类别的类条件概率密度函数之比的形式。

$$\text{即: } l_{12} = \frac{p(X|\omega_1)}{p(X|\omega_2)}$$

似然比形式的判决规则:

$$l_{12} > \theta_{12} \Rightarrow X \in \omega_1$$

$$l_{12} < \theta_{12} \Rightarrow X \in \omega_2$$

$$\text{最大后验概率判决规则中: } \theta_{12} = \frac{P(\omega_2)}{P(\omega_1)}$$

$$\text{最小风险判决规则中: } \theta_{12} = \frac{[L(\alpha_1|\omega_2) - L(\alpha_2|\omega_2)]P(\omega_2)}{[L(\alpha_2|\omega_1) - L(\alpha_1|\omega_1)]P(\omega_1)}$$

两者在似然比表现形式上面完全一样, 但是判决阈值的计算不同。若令最小风险判决准则中判决正确的风险为 0, 判决错误的代价为 1, 则二者的判决阈值计算也是一样的。所以最大后验概率判决准则是 0-1 风险判决准则。

4. 在图像识别中，假定有灌木丛和坦克两种类型，它们的先验概率分别是 0.8 和 0.2，损失函数如下表所示，其中 ω_1 和 ω_2 分别表示灌木丛和坦克， α_1 和 α_2 表示判决为灌木丛和坦克， α_3 表示拒绝判决。

现在做了三次实验，从类概率密度函数曲线上查得三个样本 X_1 ， X_2 ， X_3 的类概率密度值如下：

	ω_1	ω_2
α_1	0.5	6
α_2	2	1
α_3	1.5	1.5

$$X_1: p(X_1|\omega_1) = 0.1, p(X_1|\omega_2) = 0.7$$

$$X_2: p(X_2|\omega_1) = 0.3, p(X_2|\omega_2) = 0.45$$

$$X_3: p(X_3|\omega_1) = 0.6, p(X_3|\omega_2) = 0.5$$

- (1) 试用贝叶斯最小误判概率准则判决三个样本各属于哪一个类型。
- (2) 假定只考虑前两种判决，试用贝叶斯最小风险判决准则判决三个样本各属于哪一个类型。
- (3) 把拒绝判决考虑在内，重新考核三次实验的结果。

(1) 贝叶斯最小误判准则：

$$p(X|\omega_1)P(\omega_1) > p(X|\omega_2)P(\omega_2) \Rightarrow X \in \omega_1$$

$$p(X|\omega_1)P(\omega_1) > p(X|\omega_2)P(\omega_2) \Rightarrow X \in \omega_2$$

第一个样本：

$$p(X_1|\omega_1)P(\omega_1) = 0.1 \times 0.8 = 0.08$$

$$p(X_1|\omega_2)P(\omega_2) = 0.14$$

$$X_1 \in \omega_2$$

第二个样本：

$$p(X_2|\omega_1)P(\omega_1) = 0.24$$

$$p(X_2|\omega_2)P(\omega_2) = 0.09$$

$$X_2 \in \omega_1$$

第三个样本：

$$p(X_3|\omega_1)P(\omega_1) = 0.48$$

$$p(X_3|\omega_2)P(\omega_2) = 0.1$$

$$X_3 \in \omega_1$$

(2) 考虑前两种判决，贝叶斯最小风险判决准则

$$l_{12} > \theta_{12} \Rightarrow x \in \omega_1$$

$$l_{12} < \theta_{12} \Rightarrow x \in \omega_2$$

$$l_{12} = \frac{p(x|\omega_1)}{p(x|\omega_2)}$$

$$\theta_{12} = \frac{[L(\alpha_1|\omega_2) - L(\alpha_2|\omega_2)]P(\omega_2)}{[L(\alpha_2|\omega_1) - L(\alpha_1|\omega_1)]P(\omega_1)} = \frac{5}{6}$$

第一个样本：

$$l_{12} = \frac{1}{7} < \theta_{12} = \frac{5}{6}$$

$$X_1 \in \omega_2$$

第二个样本：

$$l_{12} = \frac{2}{3} < \theta_{12} = \frac{5}{6}$$

$$X_2 \in \omega_2$$

第三个样本：

$$l_{12} = 1.2 > \theta_{12} = \frac{5}{6}$$

$$X_3 \in \omega_1$$

(3) 考虑拒绝判决

注意：书中的推导是在假设所有的正确判决损失均为 0，所有的错误判决损失均为 λ_F ，所有的拒绝判决损失均为 λ_R 的情况下进行的。而本题正确判决损失不为 0，错误判决损失也不都相同。

$$R(\alpha_{N+1} | X) < R(\alpha_i | X), \forall i = 1, 2, \dots, N \Rightarrow X \in \alpha_{N+1}$$

$$R(\alpha_i | X) = \sum_{j=1}^N L(\alpha_i | \omega_j) P(\omega_j | X) = \sum_{j=1}^N L(\alpha_i | \omega_j) \frac{p(X | \omega_j) P(\omega_j)}{\sum_{j=1}^N p(X | \omega_j) P(\omega_j)}, i = 1, 2, 3, \dots, N+1$$

第一个样本：

$$R(\alpha_1 | X_1) = 4$$

$$R(\alpha_2 | X_1) = \frac{15}{11} \approx 1.36$$

$$R(\alpha_3 | X_1) = 1.5$$

$$X_1 \in \omega_2$$

第二个样本：

$$R(\alpha_1 | X_2) = 2$$

$$R(\alpha_2 | X_2) = \frac{19}{11} \approx 1.73$$

$$R(\alpha_3 | X_2) = 1.5$$

拒绝判决

第三个样本:

$$R(\alpha_1 | X_3) = \frac{42}{29} \approx 1.45$$

$$R(\alpha_2 | X_3) = \frac{53}{29} \approx 1.83$$

$$R(\alpha_3 | X_3) = 1.5$$

$$X_3 \in \omega_1$$

第四章习题

1、在有限样本情况下，可以使用哪些数据分组的方式？简要地概括每种方式的优缺点。

答：

可以采用两分法、留一法、分组交替法、重复使用法等数据分组方式，优缺点详见教材 P173，4.5.2 小节。

2、考虑正态分布情况下错误率的计算问题，假设，有 ω_1 、 ω_2 二分类问题，其均值向量分别为 μ_1 、 μ_2 ，协方差矩阵分别为 Σ_1 、 Σ_2 ，且 $\Sigma_1 = \Sigma_2 = \Sigma$ 。假设使用基于最大似然比的方案设计分类器，采用最小风险判决规则，请给出此情况下的总分类错误率 $P(e)$ ，说明其中各符号的含义，并分析总分类错误率和什么因素相关。

答：

根据所给条件，可得总错误率的表达式，详细推导见教材 P149-P152。

$$\begin{aligned} P(e) &= P(\omega_1)P_1(e) + P(\omega_2)P_2(e) \\ &= P(\omega_1)\Phi\left(\frac{\theta_{12} - \gamma_{ij}^2/2}{\gamma_{ij}}\right) + P(\omega_2)\left(1 - \Phi\left(\frac{\theta_{12} + \gamma_{ij}^2/2}{\gamma_{ij}}\right)\right) \end{aligned}$$

其中， $P(\omega_1)$ 、 $P(\omega_2)$ 为两个类别的先验概率， γ_{ij} 为两个类别的均值之间的马氏距离， θ_{12} 为采用最小风险判决所对应的判决阈值。

由公式可见，总的错误率与两个类别的先验概率，相应的判决阈值，类别均值之间的马氏距离有关。

第五章习题答案

1. 聚类算法可分为哪几类？影响聚类算法性能优劣的因素有哪些？

(1) 增类聚类算法（基于分裂的聚类算法）

(2) 减类聚类算法（基于合并的聚类算法）

(3) 动态聚类算法

因素：聚类中心的选择与更新、聚类策略和聚类准则的选择、控制阈值和类别数的设置等

2. 请给出最小张树聚类算法的具体过程，并分析其优缺点。

课本第 218~221 页

优点：简洁、明快，特别适用于距离较远的两个密集样本点集之间的区分。

缺点：1) 当样本集中存在噪声样本时可能造成错分；2) 对于相距较近的两个密集样本点集的区分能力较弱。

3. 有以下四个样本： $X_1 = (3, 5)^T$, $X_2 = (5, 1)^T$, $X_3 = (1, 0)^T$, $X_4 = (1, 4)^T$

初始划分为两类： $w_1: \{X_1, X_2\}$ 和 $w_2: \{X_3, X_4\}$

(1) 若将 X_2 移到 w_2 类，试计算出转移前后总的类内离散度矩阵 S_w 。

(2) 若使用 S_w 的行列式作为聚类准则，(1) 中的转移是否合适？

(3) 使用 C-均值算法的准则函数，(1) 中的转移是否合适？

(1) 若将 X_2 移到 w_2 类，试计算出转移前后的总的类内离散度矩阵 S_w 。

转移前： 样本均值 $m_1 = (4, 3)^T$, $m_2 = (1, 2)^T$

$$\text{类内离散度矩阵 } S_1 = \begin{bmatrix} 1 & -2 \\ -2 & 4 \end{bmatrix}$$

$$S_2 = \begin{bmatrix} 0 & 0 \\ 0 & 4 \end{bmatrix}$$

$$\text{总的类内离散度矩阵 } S_w = \sum_{j=1}^2 P_j S_j = \frac{1}{2} S_1 + \frac{1}{2} S_2 = \begin{bmatrix} \frac{1}{2} & -1 \\ -1 & 4 \end{bmatrix}$$

转移后： 样本均值 $m'_1 = (3, 5)^T$, $m'_2 = (\frac{7}{3}, \frac{5}{3})^T$

类内离散度矩阵 $S'_1 = 0$

$$S'_2 = \begin{bmatrix} \frac{32}{9} & -\frac{8}{9} \\ -\frac{8}{9} & \frac{26}{9} \end{bmatrix}$$

$$\text{总的类内离散度矩阵 } S'_w = \sum_{j=1}^2 P_j S_j = \frac{1}{4} S_1 + \frac{3}{4} S_2 = \begin{bmatrix} \frac{8}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{13}{6} \end{bmatrix}$$

(2) 若使用 S_w 的行列式作为聚类准则, (1)中的转移是否合适?

优化目标: S_w 的行列式越小越好

$$S_w = \begin{bmatrix} \frac{1}{2} & -1 \\ -1 & 4 \end{bmatrix} \quad S'_w = \begin{bmatrix} \frac{8}{3} & -\frac{2}{3} \\ -\frac{2}{3} & \frac{13}{6} \end{bmatrix}$$

转移前: $|S_w|=1$ 转移后: $|S'_w|=\frac{16}{3}$ 不合适!

(3) 若使用C-均值算法的准则函数, (1)中的转移是否合适?

准则函数: 误差平方和准则函数 J_e
越小越好

转移前: $J_e=18$ 转移后: $J'_e=19.33$ 不合适!

4. 有以下5个样本: $X_1 = (0, 1, 2, 1, 2, 4)^T$, $X_2 = (3, 2, 3, 1, 2, 1)^T$, $X_3 = (1, 0, 0, 0, 1, 1)^T$, $X_4 = (2, 1, 0, 2, 1, 2)^T$, $X_5 = (0, 0, 1, 0, 1, 0)^T$
请按照最小距离准则用层次聚类法进行聚类分析。

初始化: 类别计数器 $C_s = 5$

类间距离阈值 T =充分大的数

迭代计数器 $k=0$

(1) 计算各样本间的欧式距离, 构造矩阵 D

$$D^0 = \begin{bmatrix} 0 & \sqrt{20} & \sqrt{17} & \sqrt{14} & \sqrt{20} \\ \sqrt{20} & 0 & \sqrt{19} & \sqrt{14} & \sqrt{20} \\ \sqrt{17} & \sqrt{19} & 0 & \sqrt{7} & \sqrt{3} \\ \sqrt{14} & \sqrt{14} & \sqrt{7} & 0 & \sqrt{14} \\ \sqrt{20} & \sqrt{20} & \sqrt{3} & \sqrt{14} & 0 \end{bmatrix}$$

Minimize $D_{pq}: D_{35}, D_{53}$

$w_1 = \{x_1\}, w_2 = \{x_2\}, w_3 = \{x_3\}, w_4 = \{x_4\}, w_5 = \{x_5\}$

↓

$w_1 = \{x_1\}, w_2 = \{x_2\}, w_3 = \{x_3, x_5\}, w_4 = \{x_4\}$

(2) $k=1, C_S = 4$

$$D^1 = \begin{bmatrix} 0 & \sqrt{20} & \sqrt{17} & \sqrt{14} \\ \sqrt{20} & 0 & \sqrt{19} & \sqrt{14} \\ \sqrt{17} & \sqrt{19} & 0 & \sqrt{7} \\ \sqrt{14} & \sqrt{14} & \sqrt{7} & 0 \end{bmatrix}$$

$$\begin{aligned} & \text{Minimize } D_{pq}: D_{34}, D_{43} \\ & w_1 = \{x_1\}, w_2 = \{x_2\}, w_3 = \{x_3, x_5\}, w_4 = \{x_4\} \\ & \downarrow \\ & w_1 = \{x_1\}, w_2 = \{x_2\}, w_3 = \{x_3, x_5, x_4\} \end{aligned}$$

(3) $k=2, C_S = 3$

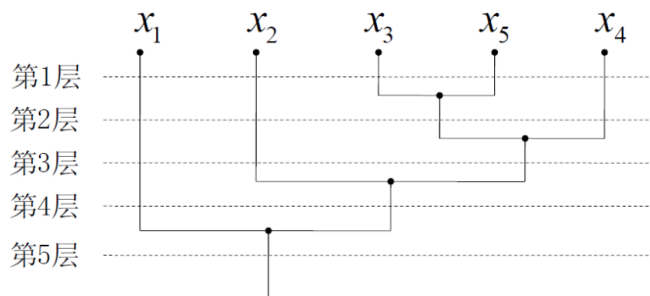
$$D^2 = \begin{bmatrix} 0 & \sqrt{20} & \sqrt{14} \\ \sqrt{20} & 0 & \sqrt{14} \\ \sqrt{14} & \sqrt{14} & 0 \end{bmatrix}$$

$$\begin{aligned} & \text{Minimize } D_{pq}: D_{23}, D_{32} \\ & w_1 = \{x_1\}, w_2 = \{x_2\}, w_3 = \{x_3, x_5, x_4\} \\ & \downarrow \\ & w_1 = \{x_1\}, w_2 = \{x_2, x_3, x_5, x_4\} \\ & \text{Minimize } D_{pq}: D_{13}, D_{31} \\ & w_1 = \{x_1\}, w_2 = \{x_2\}, w_3 = \{x_3, x_5, x_4\} \\ & \downarrow \\ & w_1 = \{x_1, x_3, x_5, x_4\}, w_2 = \{x_2\} \end{aligned}$$

(4) $k=3, C_S = 2$

$$D^3 = \begin{bmatrix} 0 & \sqrt{18} \\ \sqrt{18} & 0 \end{bmatrix}$$

$$\begin{aligned} & w_1 = \{x_1\}, w_2 = \{x_2, x_3, x_5, x_4\} \\ & \downarrow \\ & w_1 = \{x_1, x_2, x_3, x_5, x_4\} \end{aligned}$$



5. 已知一个样本集合: $\{(0, 0)^T, (1, 1)^T, (2, 1)^T, (4, 4)^T, (5, 3)^T, (5, 4)^T, (6, 3)^T, (6, 5)^T\}$
试用近邻函数法进行聚类分析。

(1) 计算距离矩阵D

$$D = \begin{bmatrix} 0 & \sqrt{2} & \sqrt{5} & \sqrt{32} & \sqrt{34} & \sqrt{41} & \sqrt{45} & \sqrt{61} \\ \sqrt{2} & 0 & 1 & \sqrt{18} & \sqrt{20} & 5 & \sqrt{29} & \sqrt{41} \\ \sqrt{5} & 1 & 0 & \sqrt{13} & \sqrt{13} & \sqrt{18} & \sqrt{20} & \sqrt{32} \\ \sqrt{32} & \sqrt{18} & \sqrt{13} & 0 & \sqrt{2} & 1 & \sqrt{5} & \sqrt{5} \\ \sqrt{34} & \sqrt{20} & \sqrt{13} & \sqrt{2} & 0 & 1 & 1 & \sqrt{5} \\ \sqrt{41} & 5 & \sqrt{18} & 1 & 1 & 0 & \sqrt{2} & \sqrt{2} \\ \sqrt{45} & \sqrt{29} & \sqrt{20} & \sqrt{5} & 1 & \sqrt{2} & 0 & 2 \\ \sqrt{61} & \sqrt{41} & \sqrt{32} & \sqrt{5} & \sqrt{5} & \sqrt{2} & 2 & 0 \end{bmatrix}$$

(2) 对矩阵的每一列进行排序，得到近邻系数矩阵M

m_{ij} : X_i 对 X_j 的近邻系数 (X_i 是 X_j 的第 m_{ij} 个近邻)

$$M = \begin{bmatrix} 0 & 2 & 2 & 7 & 7 & 7 & 7 \\ 1 & 0 & 1 & 6 & 6 & 6 & 6 \\ 2 & 1 & 0 & 5 & 5 & 5 & 5 \\ 3 & 3 & 3 & 0 & 3 & 1 & 4 \\ 4 & 4 & 3 & 2 & 0 & 1 & 1 \\ 5 & 5 & 5 & 1 & 1 & 0 & 2 \\ 6 & 6 & 6 & 3 & 1 & 3 & 0 \\ 7 & 7 & 7 & 3 & 4 & 3 & 3 \end{bmatrix}$$

(3) 计算近邻函数值矩阵 $A = [\alpha_{ij}]$, $\alpha_{ij} = m_{ij} + m_{ji} - 2$

$$A = \begin{bmatrix} 16 & 1 & 2 & 8 & 9 & 10 & 11 & 12 \\ 1 & 16 & 0 & 7 & 8 & 9 & 10 & 11 \\ 2 & 0 & 16 & 6 & 6 & 8 & 9 & 10 \\ 8 & 7 & 6 & 16 & 3 & 0 & 5 & 4 \\ 9 & 8 & 6 & 3 & 16 & 0 & 0 & 5 \\ 10 & 9 & 8 & 0 & 0 & 16 & 3 & 2 \\ 11 & 10 & 9 & 5 & 0 & 3 & 16 & 3 \\ 12 & 11 & 10 & 4 & 5 & 2 & 3 & 16 \end{bmatrix}$$

$$\omega_1 = \{x_1, x_2, x_3\}$$

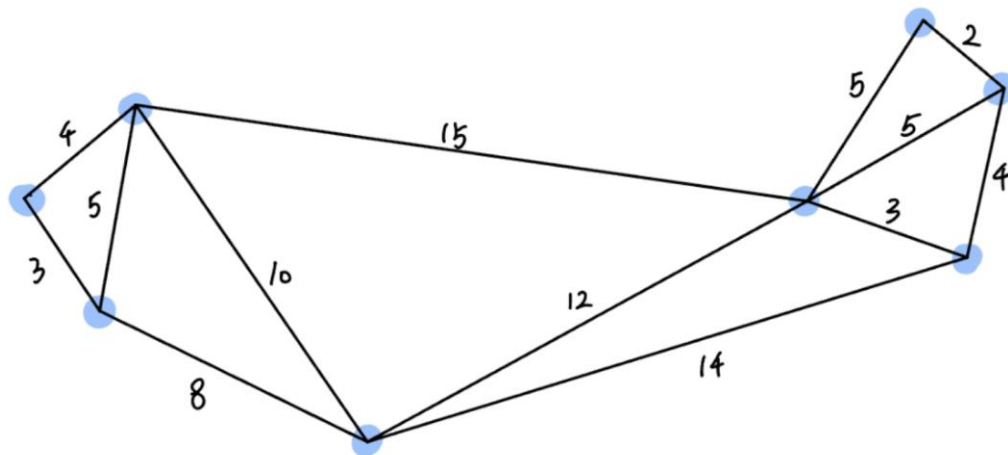
$$\omega_2 = \{x_4, x_5, x_6, x_7, x_8\}$$

$$\gamma_{12} = 6, \max \gamma_1 = 1, \max \gamma_2 = 2$$

$$\gamma_{12} > \gamma_1, \gamma_{12} > \gamma_2$$

分类结束

6. 假设某样本集合的加权图如下表示，请给出对应的最小张树表示，并通过最小张树，将样本划为两类。(要求给出中间结果)。



第六章习题答案

1. 简述文法的分类以及相应的定义.

答: 文法分为四类:0 型文法、1 型文法、2 型文法和 3 型文法。

0 型文法: 也称为无约束文法或短语结构文法,其产生式具有

$$\alpha \rightarrow \beta$$

的形式. 其中, $\alpha \in \Sigma^+$ 和 $\beta \in \Sigma^*$. 这类文法对产生式没有任何限制.

1 型文法: 也称为上下文有关文法,其产生式具有

$$\alpha_1 A \alpha_2 \rightarrow \alpha_1 \beta \alpha_2$$

的形式. 其中, $\alpha_1, \alpha_2 \in \Sigma^*$, $\beta \in \Sigma^+$ 以及 $A \in N$.

2 型文法: 也称为上下文无关文法, 其产生式具有

$$A \rightarrow \beta$$

的形式. 其中, $A \in N, \beta \in \Sigma^+$.

3 型文法: 也称为有限状态文法或正则文法, 其产生式具有

$$A \rightarrow aB \quad \text{或} \quad A \rightarrow b$$

的形式. 其中, $A, B \in N, a, b \in T$.

2. 考虑文法 $G = (N, T, P, S)$, 其中 $N = \{S, A, B\}$, $T = \{a, b, c\}$, 以及 P :

(1) $S \rightarrow aAc$ (2) $A \rightarrow cBb$ (3) $B \rightarrow aBa$ (4) $B \rightarrow b$

1) 文法 G 是什么型文法?

2) 文法 G 可以生成的语言是什么? (要求详细推导过程)

解:

1) 文法 G 是二型文法.

2) 文法 G 可以生成的语言是 $L(G) = \{aca^nba^nbc | n = 0, 1, 2, \dots\}$

推导过程如下:

(1)	(2)	(4)			
$S \Rightarrow aAc$	$\Rightarrow acBbc$	$\Rightarrow acbbc$			
G	G	G			
(1)	(2)	(3)	(4)		
$S \Rightarrow aAc$	$\Rightarrow acBbc$	$\Rightarrow acaBabc$	$\Rightarrow acababc$		
G	G	G	G		
(1)	(2)	(3)	(3)	(4)	
$S \Rightarrow aAc$	$\Rightarrow acBbc$	$\Rightarrow acaBabc$	$\Rightarrow acaaBaabc$	$\Rightarrow acaabaabc$	
G	G	G	G	G	
(1)	(2)	(3)	(3)	(3)	(4)
$S \Rightarrow aAc$	$\Rightarrow acBbc$	$\Rightarrow acaBabc$	$\Rightarrow acaaBaabc$	$\Rightarrow acaaaBaaabc$	$\Rightarrow acaaabaaabc$
G	G	G	G	G	G
.....					

所以, 由文法 G 生成的语言是

$L(G) = \{acbbc, acababc, aca^2ba^2bc, aca^3ba^3bc, \dots\} = \{aca^nba^nbc | n = 0, 1, 2, \dots\}$

3. 已知一个非确定的有限状态自动机 $A = (Q, \Sigma, \delta, q_0, F)$, 其中:

$Q = \{q_0, q_1, q_2\}$, $\Sigma = \{0, 1\}$, $F = \{q_2\}$, 以及 δ :

$$(1) \delta(q_0, 0) = \{q_0, q_1\} \quad (2) \delta(q_0, 1) = \{q_0, q_2\}$$

$$(3) \delta(q_1, 0) = \{q_2\} \quad (4) \delta(q_1, 1) = \{q_1\}$$

$$(5) \delta(q_2, 0) = \{q_2\} \quad (6) \delta(q_2, 1) = \{q_2\}$$

1) 画出该非确定有限状态自动机的状态转移图.

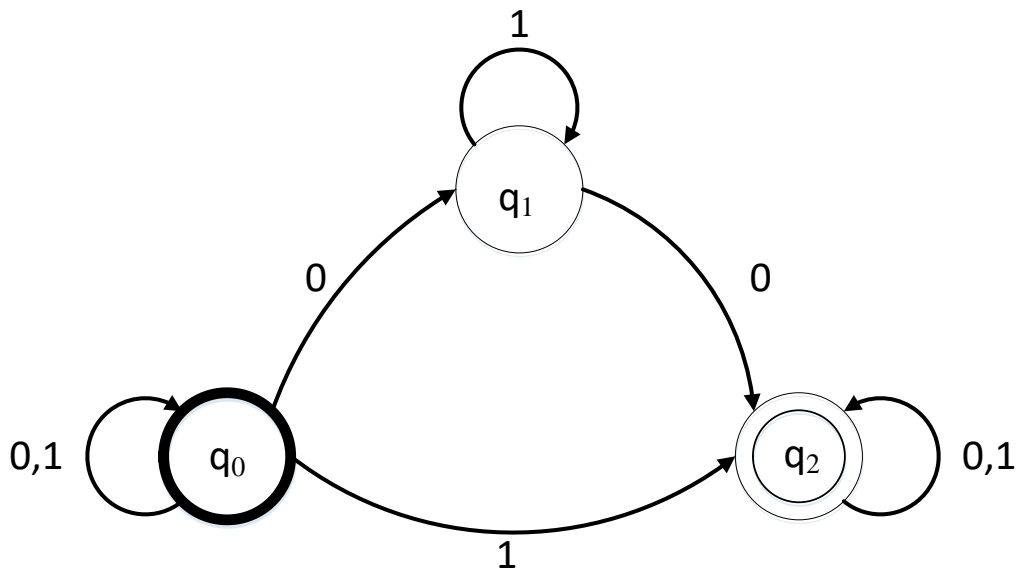
2) 构造对应的确定的有限状态自动机, 并给出状态转移图.

解:

1) 状态转移表:

符号 状态	0	1
$\cdot q_0$	$\{ \cdot q_0, q_1 \}$	$\{ \cdot q_0, q_2 \}$
q_1	$q_2 \cdot$	q_1
$q_2 \cdot$	$q_2 \cdot$	$q_2 \cdot$

状态转移图:

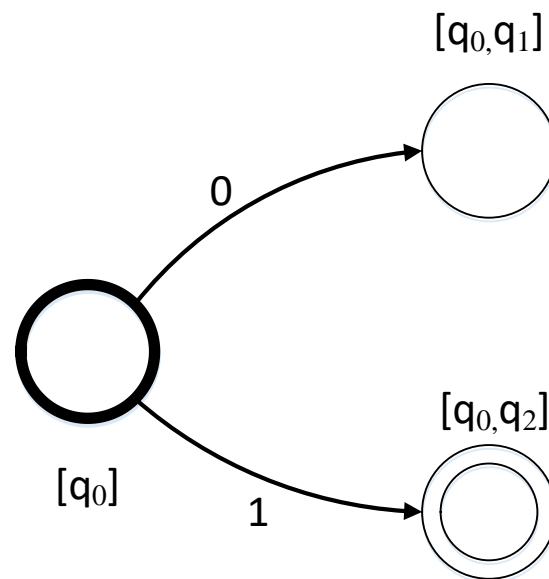


2) 设所求的确定有限状态自动机 $A' = (Q', \Sigma, \delta', q'_0, F')$, 其可按照以下步骤获得:

① A' 的初态为 $[q_0]$.



② 考虑初始状态 $[q_0]$ 在输入符号为 0, 1 情况下的状态转移情况. 因在 δ 中有 $\delta(q_0, 0) = \{q_0, q_1\}$, 而 A' 中无对应的转移状态, 故新增状态 $[q_0, q_1]$, 又因在 δ 中有 $\delta(q_0, 1) = \{q_0, q_2\}$, 而在 A' 中无对应的转移状态, 故增加状态 $[q_0, q_2]$, 由于 $\{q_0, q_2\}$ 中含有 A 的终结状态 q_2 , 故 $[q_0, q_2]$ 为 A' 的终结状态.

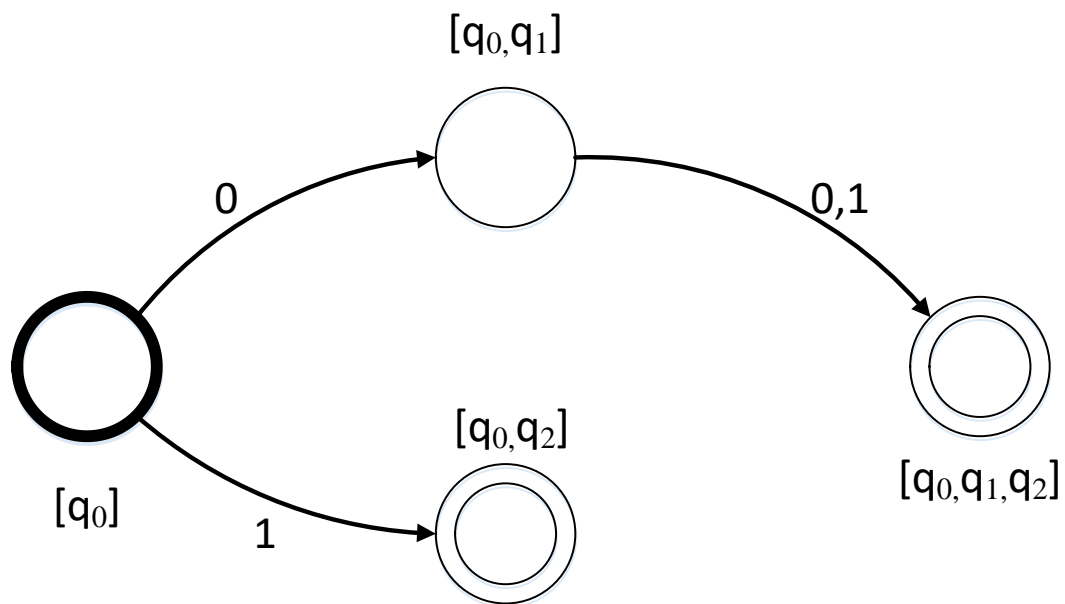


③ 考虑状态 $[q_0, q_1]$ 在输入符号为 0, 1 情况下的状态转移情况. 在 δ 中有

$$\delta(\{q_0, q_1\}, 0) = \{q_0, q_1, q_2\}$$

$$\delta(\{q_0, q_1\}, 1) = \{q_0, q_1, q_2\}$$

而在 A' 中无对应的转移状态, 故新增状态 $[q_0, q_1, q_2]$, 由于 $\{q_0, q_1, q_2\}$ 中含有 A 的终结状态 q_2 , 故 $[q_0, q_1, q_2]$ 为 A' 的终结状态.

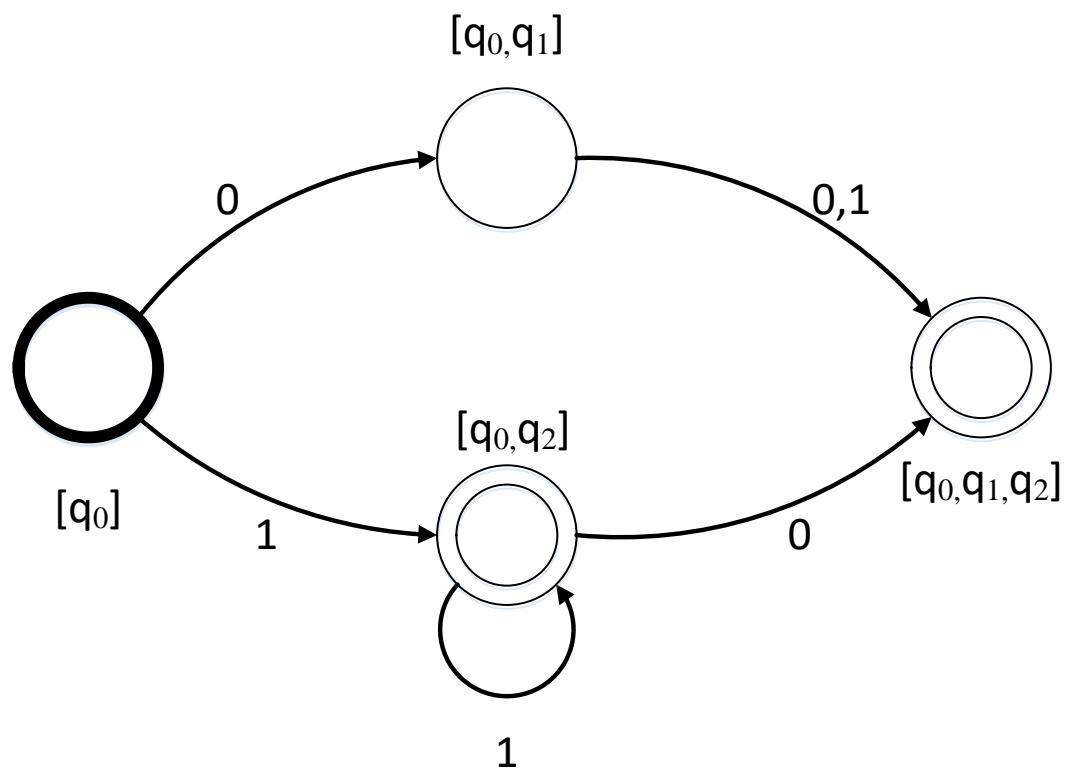


④ 考虑 $[q_0, q_2]$ 在输入符号为 0, 1 情况下的状态转移情况.

$$\delta(\{q_0, q_2\}, 0) = \{q_0, q_1, q_2\}$$

$$\delta(\{q_0, q_2\}, 1) = \{q_0, q_2\}$$

而在 A 中已有状态 $[q_0, q_2]$ 和 $[q_0, q_1, q_2]$, 故利用已有的状态.



⑤ 考虑 $[q_0, q_1, q_2]$ 在输入符号为 0, 1 情况下的状态转移情况.

$$\delta(\{q_0, q_1, q_2\}, 0) = \{q_0, q_1, q_2\}$$

$$\delta(\{q_0, q_1, q_2\}, 1) = \{q_0, q_1, q_2\}$$

而在 A 中已有状态 $[q_0, q_1, q_2]$, 故利用已有的状态.

至此, A 中所有的状态已被遍历, 故所得状态转移图即为所求状态转移图.

