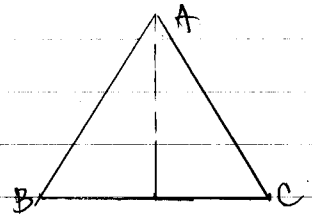


习题课 (一)

1 (第6章) 课本 6-6. $d_{AB}^2 = d_{AC}^2 = 1.1$, $d_{BC}^2 = 1.0$.

试用例题说明重心法不具有单调性



2	总统	出生地	是否竞选?	党派	是否经力	曾任副总统
(第6章)	1. 里根	中西部	Y	共和	N	N
	2. 卡特	南方	Y	民主	N	N
	3. 福特	中西部	N	共和	Y	Y
	4. 尼克松	西部	Y	共和	Y	Y
	5. 约翰逊	南方	N	民主	Y	Y
	6. 肯尼迪	东部	Y	民主	Y	N

(a) 引进适当的二值变量, 对各对总统计算匹配系数 ($r_{ij} = \frac{n_{01} + m_{11}}{p}$) 阵(b) 对总统计算匹配系数 ($r'_{ij} = \frac{2(m_{01} + m_{11})}{2(m_{01} + m_{11}) + m_{21}}$) 阵, 并验证这两种系数定义有单调关系 (i.e. $r_{ij} > r_{kl} \Leftrightarrow r'_{ij} > r'_{kl}$)

(c) 证明 (a) (b) 两种匹配系数定义有单调关系

3 (第7章) 课本 7-9.

4 (第7章) 课本 7-10.

(第7章)

5. 令 $z_0, z_1, \dots, z_d \sim i.i.d. (0, \sigma^2)$, 令 $x_i = z_0 + z_i$ ($i=1, 2, \dots, d$) $\mathbf{x}' = (x_1, \dots, x_d)'$ ① 证明存在 $\mathbf{x}$ 的主成分 $\propto z_i$,

② 求 ① 中主成分的方差

特殊结构
主成分
阵的正交

6. 五种股票间的相关系数阵.

(第6章)

$$\begin{pmatrix} 1 & & & & \\ 0.63 & 1 & & & \\ 0.51 & 0.57 & 1 & & \\ 0.12 & 0.32 & 0.18 & 1 & \\ 0.16 & 0.21 & 0.15 & 0.68 & 1 \end{pmatrix}$$

将样本相关系数用作相似性度量, 用类平均法作系统聚类分析.

7. 一项调查提供了威斯康星州某地区的五项社会经济变量:
第7章) X_1 : 总人口份 X_2 : 职业等级(%) X_3 : 16岁以上人员失业率(%)
 X_4 : 政府雇佣(%) X_5 : 家庭收入中位数(10万元). 从61个小区中得到的数据综合成计量:

$$\bar{x} = (4.47 \quad 3.96 \quad 71.42 \quad 26.91 \quad 1.64)'$$

样本协方差阵及特征及特征向量

$$\begin{aligned} \lambda_1 &= 107.02 & \vec{e}_1' &= (-0.039 \quad 0.105 \quad -0.492 \quad 0.863 \quad 0.009) \\ \lambda_2 &= 39.67 & \vec{e}_2' &= (0.071 \quad 0.130 \quad 0.864 \quad 0.980 \quad 0.015) \\ \lambda_3 &= 8.37 & \vec{e}_3' &= (0.188 \quad -0.961 \quad 0.046 \quad -0.153 \quad -0.125) \\ \lambda_4 &= 2.87 & \vec{e}_4' &= (0.977 \quad 0.171 \quad -0.091 \quad -0.030 \quad 0.082) \\ \lambda_5 &= 0.15 & \vec{e}_5' &= (-0.058 \quad -0.139 \quad 0.005 \quad 0.007 \quad 0.989) \end{aligned}$$

① 求因子载荷阵

② 求累计贡献率

③ 根据①②确定主成分个数, 并求所确定的主成分对 $X_i (i=1, \dots, 5)$ 的贡献率

$$S = \begin{pmatrix} 339.1 & -7.102 & 4.306 & -2.078 & 0.007 \\ & 9.673 & -1.513 & 10.953 & 1.203 \\ & & 55.626 & -28.937 & -0.044 \\ & & & 89.067 & 0.957 \\ & & & & 0.319 \end{pmatrix}$$

④ 简单描述③中所确定的主成分的实际意义.

8 ① 课本 8-1, 8-2.

(2) 计算 X_i 的公共度, 并解释这些量.

(3) 哪个变量会在公共因子的命名上有最大的贡献, 计算相应的统计量来证明你的观点.

(4) 设 m 个因子模型, 计算因子载荷阵 A 的主因子.

9. 鸡的骨骼测量数据: X_1 : 颅骨长度 X_2 : 颅骨幅宽
(第8章) X_3 : 股骨长 X_4 : 胫骨长
 X_5 : 跗骨长 X_6 : 尺骨长

样本相关阵:

$$\begin{pmatrix} 1.000 & & & & & \\ 0.505 & 1.000 & & & & \\ 0.569 & 0.422 & 1.000 & & & \\ 0.602 & 0.467 & 0.925 & 1.000 & & \\ 0.621 & 0.482 & 0.877 & 0.874 & 1.000 & \\ 0.603 & 0.450 & 0.878 & 0.894 & 0.937 & 1.000 \end{pmatrix}$$

以下因子载荷由极大似然法求出

X_i	因子载荷估计		因子载荷的最大方差旋转	
	F_1	F_2	F_1	F_2
X_1	0.200	0.602	0.411	0.484
X_2	0.154	0.467	0.319	0.375
X_3	0.143	0.926	0.717	0.603
X_4	0.000	1.000	0.855	0.519
X_5	0.476	0.874	0.499	0.861
X_6	0.327	0.894	0.594	0.744

1) 用未旋转的因子载荷估计求以下各项的极大似然估计:
(a) 特殊方差 (b) 共同度 (c) 各因子对总差的贡献率
(d) 残差阵

(2) 用未旋转和旋转后以结果计算最大方差准则的值.

(3) 随机选取一新样本测得 $\bar{X} = (39.1, 39.3, 75.7, 115, 73.4, 69.1)'$,
用加权最小二乘法计算其因子得分.

10. 课本 8-4