应用多元统计分析作业

朱强强
17064001
应统1701

7.1 设$\boldsymbol{x}=(x_1,x_2,\cdots,x_p)'$的协方差矩阵为

$$\Sigma=\begin{pmatrix}\sigma_{11} & & & 0\\ & \sigma_{22} & & \\ & & \ddots & \\ 0 & & & \sigma_{pp}\end{pmatrix}$$

其中$\sigma_{11}\geq\sigma_{22}\geq\cdots\geq\sigma_{pp}$，试求$\boldsymbol{x}$的主成分及主成分具有的特征值。该题说明了什么？

第i个主成分为$y_i=x_i$，具有方差σ_{ii}。说明当原始变量互不相关时，主成分为原始变量。

7.2 试证多元正态变量的主成分仍为正态变量且相互独立。

$\boldsymbol{y}=\boldsymbol{T}'\boldsymbol{x}$，因为$\boldsymbol{x}$服从多元正态分布，而$\boldsymbol{T}$是一个正交矩阵，多元正态分布的线性变换仍为正态分布，所以$\boldsymbol{y}$仍为服从正态分布。所以主成分$y_1,y_2,\cdots,y_n$皆服从正态分布，又它们互不相关，所以相互独立。

7.3 二维随机变量$\boldsymbol{x}=(x_1,x_2)'$的相关矩阵总能表示为

$$\boldsymbol{R}=\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}$$

故当$\rho\neq 0$时从$\boldsymbol{R}$出发的$\boldsymbol{x}$的主成分及其贡献率应有统一的表达式，试求之。主成分所在方向与ρ有关吗？

经计算，$\boldsymbol{R}$的特征值及特征向量为

$$\lambda_1=1+|\rho| \qquad \lambda_2=1-|\rho|$$

当$\rho>0$时，$y_1=\dfrac{1}{\sqrt{2}}x_1^*-\dfrac{1}{\sqrt{2}}x_2^*,y_2=\dfrac{1}{\sqrt{2}}x_1^*+\dfrac{1}{\sqrt{2}}x_2^*$

当$\rho<0$时，$y_1=\dfrac{1}{\sqrt{2}}x_1^*+\dfrac{1}{\sqrt{2}}x_2^*,y_2=\dfrac{1}{\sqrt{2}}x_1^*-\dfrac{1}{\sqrt{2}}x_2^*$

所以主成分所在方向与ρ的绝对值大小无关，只与ρ的取值有关。

补充作业 P205例7.2.1 求y_3对每个原始变量的贡献率，要求列出具体公式

$$\rho_{1\cdot 3} = 0.924^2 \times 0.17 = 0.145$$

$$\rho_{2\cdot 3} = 0.383^2 \times 0.17 = 0.025$$

$$\rho_{3\cdot 3} = 0$$