《多元统计分析》课后作业

姓名：李倩倩

学号：2024017349

班级：统计 24-1 班

中国石油大学（北京）克拉玛依校区文理学院数学与统计系

Friday 27th March, 2026

作业要求

1. 可以和其他同学讨论作业当中的问题，但应当自己独立完成作业
2. 计算、证明等要有过程，要有主要步骤的说明
3. 请将计算、绘图所用的 R 代码以及生成的结果和图像一并添加在作业文件当中
4. 请使用 LaTeX编辑并生成 PDF 格式的文件，第X周作业文件命名方式：学号-姓名-X.pdf
5. 评分标准：每一问得分 $\in \{2, 1, 0\}$
   - 2: 按时完成并上交作业，且答案基本正确
   - 1: 按时完成并上交作业，且答案部分正确
   - 0: 答案完全错误，或者迟交作业(规定时间72小时之后)
6. 请将完成的 PDF 格式的作业文件发送至邮箱：xiaolei@cup.edu.cn
7. 每位同学可以有一次迟交作业的机会，但不得晚于规定时间三日之后
8. 第2周作业截止时间：2026年3月27日24:00

目录

Chapter 1

第2周作业

第2周作业完成时间： Friday 27th March, 2026 21:21

1. 设 $\boldsymbol{a}$ 是一个 $(p \times 1)$ 向量，$\mathcal{A} = \mathcal{A}^{\mathrm{T}}$ 是一个对称的 $(p \times p)$ 矩阵.

   (a) [2 分] 证明

$$\frac{\partial \boldsymbol{a}^{\mathrm{T}}\boldsymbol{x}}{\partial \boldsymbol{x}} = \frac{\partial \boldsymbol{x}^{\mathrm{T}}\boldsymbol{a}}{\partial \boldsymbol{x}} = \boldsymbol{a} \tag{1.1}$$

【证明】

证明. 设 $\boldsymbol{a} = (a_1, a_2, \ldots, a_p)^{\mathrm{T}}$，$\boldsymbol{x} = (x_1, x_2, \ldots, x_p)^{\mathrm{T}}$。则

$$\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x} = \sum_{i=1}^{p} a_i x_i = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{a}.$$

由向量对向量的导数定义，

$$\frac{\partial(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x})}{\partial \boldsymbol{x}} = \left(\frac{\partial}{\partial x_1}(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x}),\ \frac{\partial}{\partial x_2}(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x}),\ \ldots,\ \frac{\partial}{\partial x_p}(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x})\right)^{\mathrm{T}}.$$

而

$$\frac{\partial}{\partial x_i}(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x}) = \frac{\partial}{\partial x_i}\sum_{j=1}^{p} a_j x_j = a_i,$$

因此

$$\frac{\partial(\boldsymbol{a}^{\mathrm{T}}\boldsymbol{x})}{\partial \boldsymbol{x}} = (a_1, a_2, \ldots, a_p)^{\mathrm{T}} = \boldsymbol{a}.$$

同理，$\frac{\partial(\boldsymbol{x}^{\mathrm{T}}\boldsymbol{a})}{\partial \boldsymbol{x}} = \boldsymbol{a}$，故等式成立。 □

   (b) [2 分] 证明

$$\frac{\partial \boldsymbol{x}^{\mathrm{T}}\mathcal{A}\boldsymbol{x}}{\partial \boldsymbol{x}} = 2\mathcal{A}\boldsymbol{x} \tag{1.2}$$

【证明】

证明. 设 $\mathcal{A}=(a_{ij})_{p\times p}$ 因为该矩阵为对称矩阵，故满足 $a_{ij}=a_{ji}$。二次型可展开为

$$\boldsymbol{x}^{\mathrm{T}}\mathcal{A}\boldsymbol{x}=\sum_{i=1}^{p}\sum_{j=1}^{p}a_{ij}x_ix_j.$$

对 $\boldsymbol{x}$ 求导，即逐分量求导：

$$\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^{\mathrm{T}}\mathcal{A}\boldsymbol{x})=\left(\frac{\partial}{\partial x_1},\frac{\partial}{\partial x_2},\ldots,\frac{\partial}{\partial x_p}\right)^{\mathrm{T}}\left(\sum_{i=1}^{p}\sum_{j=1}^{p}a_{ij}x_ix_j\right).$$

对于第 k 个分量：

$$\frac{\partial}{\partial x_k}\left(\sum_{i=1}^{p}\sum_{j=1}^{p}a_{ij}x_ix_j\right)=\sum_{i=1}^{p}\sum_{j=1}^{p}a_{ij}\frac{\partial}{\partial x_k}(x_ix_j).$$

由于

$$\frac{\partial}{\partial x_k}(x_ix_j)=\delta_{ik}x_j+\delta_{jk}x_i,$$

其中 δ_{ik} 是 Kronecker 符号，当i=k时为1，否则为0，代入得

$$\frac{\partial}{\partial x_k}(\boldsymbol{x}^{\mathrm{T}}\mathcal{A}\boldsymbol{x})=\sum_{j=1}^{p}a_{kj}x_j+\sum_{i=1}^{p}a_{ik}x_i.$$

由对称性可知 $a_{ik}=a_{ki}$，故第二项可改写为 $\sum_{i=1}^{p}a_{ki}x_i$，因此

$$\frac{\partial}{\partial x_k}(\boldsymbol{x}^{\mathrm{T}}\mathcal{A}\boldsymbol{x})=\sum_{j=1}^{p}a_{kj}x_j+\sum_{i=1}^{p}a_{ki}x_i=2\sum_{j=1}^{p}a_{kj}x_j=2(\mathcal{A}\boldsymbol{x})_k.$$

将分量结果组装成向量即可得

$$\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^{\mathrm{T}}\mathcal{A}\boldsymbol{x})=2\mathcal{A}\boldsymbol{x}.$$

□

(c) [2 分] 证明二次型 $Q(\boldsymbol{x})=\boldsymbol{x}^{\mathrm{T}}\mathcal{A}\boldsymbol{x}$ 的 Hessian 矩阵为

$$\frac{\partial^2\boldsymbol{x}^{\mathrm{T}}\mathcal{A}\boldsymbol{x}}{\partial\boldsymbol{x}\partial\boldsymbol{x}^{\mathrm{T}}}=2\mathcal{A} \tag{1.3}$$

【证明】

证明. 由上一问结论，梯度为

$$\frac{\partial}{\partial \boldsymbol{x}}(\boldsymbol{x}^{\mathrm{T}}\mathcal{A}\boldsymbol{x})=2\mathcal{A}\boldsymbol{x}.$$

Hessian 矩阵是梯度向量对 $\boldsymbol{x}^{\mathrm{T}}$ 的导数，即

$$\frac{\partial^2}{\partial\boldsymbol{x}\partial\boldsymbol{x}^{\mathrm{T}}}(\boldsymbol{x}^{\mathrm{T}}\mathcal{A}\boldsymbol{x})=\frac{\partial}{\partial\boldsymbol{x}^{\mathrm{T}}}(2\mathcal{A}\boldsymbol{x}).$$

由于 $2\mathcal{A}\boldsymbol{x}$ 是 $\boldsymbol{x}$ 的线性函数，其导数为常数矩阵 $2\mathcal{A}$。具体地，对于线性函数 $\boldsymbol{y}=\boldsymbol{M}\boldsymbol{x}$，有

$\frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}^{\mathrm{T}}} = \boldsymbol{M}$。因此

$$\frac{\partial}{\partial \boldsymbol{x}^{\mathrm{T}}}(2\mathcal{A}\boldsymbol{x}) = 2\mathcal{A}.$$

于是

$$\frac{\partial^2 \boldsymbol{x}^{\mathrm{T}}\mathcal{A}\boldsymbol{x}}{\partial \boldsymbol{x}\partial \boldsymbol{x}^{\mathrm{T}}} = 2\mathcal{A}.$$

□

2. [2 分] 证明一个投影矩阵的特征值仅取值于集合 $\{0, 1\}$ 中.

【证明】

证明. 设 $\boldsymbol{P}$ 是一个投影矩阵，即为幂等矩阵，满足 $\boldsymbol{P}^2 = \boldsymbol{P}$。设 λ 是 $\boldsymbol{P}$ 的任一特征值，$\boldsymbol{x} \neq \boldsymbol{0}$ 为对应的特征向量，则

$$\boldsymbol{P}\boldsymbol{x} = \lambda \boldsymbol{x}.$$

两边左乘 $\boldsymbol{P}$ 得

$$\boldsymbol{P}^2\boldsymbol{x} = \lambda \boldsymbol{P}\boldsymbol{x} = \lambda^2 \boldsymbol{x}.$$

但由 $\boldsymbol{P}^2 = \boldsymbol{P}$，有 $\boldsymbol{P}^2\boldsymbol{x} = \boldsymbol{P}\boldsymbol{x} = \lambda\boldsymbol{x}$。因此

$$\lambda\boldsymbol{x} = \lambda^2\boldsymbol{x} \quad \Longrightarrow \quad (\lambda^2 - \lambda)\boldsymbol{x} = \boldsymbol{0}.$$

由于 $\boldsymbol{x} \neq \boldsymbol{0}$，得 $\lambda^2 - \lambda = 0$，即 $\lambda(\lambda - 1) = 0$，故 $\lambda = 0$ 或 $\lambda = 1$。因此，投影矩阵的特征值只能取自集合 $\{0, 1\}$。 □

3. [2 分] 作度量矩阵为 $\mathcal{A} = \Sigma^{-1}$ 的某个等距椭球体的图形，其中

$$\Sigma = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \tag{1.4}$$

【解】

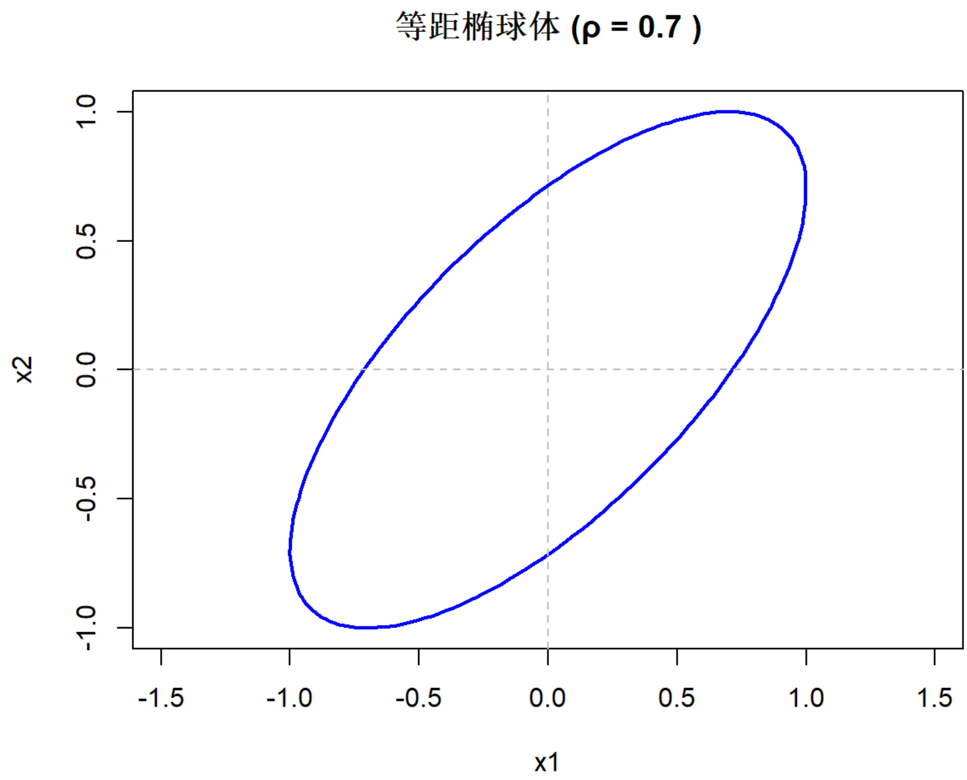

图 1.1: 等距椭圆体（当取 $\rho = 0.7$ 时）

访问 http://127.0.0.1:7810/ 可手动查看 ρ取值不同时等距椭球体的动态变化

```
  # 安装并加载 shiny
if (!require(shiny)) install.packages("shiny")
library(shiny)

ui <- fluidPage(
  titlePanel("等距椭球体: 度量矩阵 A = Σ□¹"),
  sidebarLayout(
    sidebarPanel(
      sliderInput("rho", "相关系数 ρ :",
                  min = -0.99, max = 0.99, value = 0.7, step = 0.01)
    ),
    mainPanel(
      plotOutput("ellipsePlot")
    )
  )
)
```

```
server <- function(input, output) {
  output$ellipsePlot <- renderPlot({
    rho <- input$rho
    Sigma <- matrix(c(1, rho, rho, 1), nrow = 2)
    A <- solve(Sigma)          # 度量矩阵

    # 特征分解得到椭圆的方向和半轴长度
    eig <- eigen(A)
    # 半轴长度 = 1 / sqrt(λ), 其中 λ 是 A 的特征值
    half_len <- 1 / sqrt(eig$values)

    # 参数化椭圆点（未旋转）
    theta <- seq(0, 2*pi, length.out = 200)
    x0 <- half_len[1] * cos(theta)
    y0 <- half_len[2] * sin(theta)

    # 旋转到特征向量方向
    R <- eig$vectors
    pts <- t(R %*% rbind(x0, y0))

    # 构造标题，避免 bquote 中的语法冲突
    title_text <- paste0("等距椭球体: x□ A x = 1,  ρ = ", round(rho, 3))

    # 绘图
    plot(pts, type = "l", asp = 1, col = "blue", lwd = 2,
         xlab = expression(x[1]), ylab = expression(x[2]),
         main = title_text)
    abline(h = 0, v = 0, lty = 2, col = "gray")
    grid()
  })
}

shinyApp(ui, server)
```

4. 对于课堂中讨论过的汽车数据集，

(a) [2 分] 计算变量 X_2 = miles per gallon 与 X_8 = weight 的协方差.

【解】变量 X_2 与 X_8，它们之间的协方差为

$$\sigma_{X_2X_8} = \mathrm{Cov}(X_2, X_8) = \mathbb{E}\big[(X_2 - \mathbb{E}(X_2))(X_8 - \mathbb{E}(X_8))\big] = \mathbb{E}(X_2X_8) - \mathbb{E}(X_2)\mathbb{E}(X_8).$$

用R语言计算

```
cov(auto$MPG, auto$Weight)
```

算出来X_2 = miles per gallon 与 X_8 = weight 的协方差为-3732.025

(b) [2 分] 你期待协方差的符号是正还是负，为什么?

【解】我期待的协方差为负。因为汽车油耗（每加仑英里数）与车重通常呈负相关：车重越大，燃油效率越低（mpg 越小）。因此，当车重增加时，mpg 倾向于减少，协方差为负。实际计算结果也验证了这一预期。

5. 一位纺织店经理研究“经典蓝色”套头衫在 10 个不同时期的销售情况. 他调查了销量 (X_1)；价格的变化 (X_2)，单位：欧元；当地报纸的广告费用 (X_3)，单位：欧元；以及是否有促销员 (X_4)，促销员的时长，单位：小时. 所得观测数据矩阵如下：

$$\mathcal{X} = \begin{pmatrix} 230 & 125 & 200 & 109 \\ 181 & 99 & 55 & 107 \\ 165 & 97 & 105 & 98 \\ 150 & 115 & 85 & 71 \\ 97 & 120 & 0 & 82 \\ 192 & 100 & 150 & 103 \\ 181 & 80 & 85 & 111 \\ 189 & 90 & 120 & 93 \\ 172 & 95 & 110 & 86 \\ 170 & 125 & 130 & 78 \end{pmatrix} \tag{1.5}$$

(a) [2 分] 计算 $\mathcal{X}$ 的样本相关矩阵.

【解】 设 $n = 10$，变量依次为 X_1（销量），X_2（价格变化），X_3（广告费用），X_4（促销员时长）。先计算各变量的样本均值：

$$\bar{x}_1 = \frac{230 + 181 + 165 + 150 + 97 + 192 + 181 + 189 + 172 + 170}{10} = 172.7,$$
$$\bar{x}_2 = \frac{125 + 99 + 97 + 115 + 120 + 100 + 80 + 90 + 95 + 125}{10} = 104.6,$$
$$\bar{x}_3 = \frac{200 + 55 + 105 + 85 + 0 + 150 + 85 + 120 + 110 + 130}{10} = 104.0,$$
$$\bar{x}_4 = \frac{109 + 107 + 98 + 71 + 82 + 103 + 111 + 93 + 86 + 78}{10} = 93.8.$$

然后计算离差平方和与交叉积（$n-1=9$），得样本协方差矩阵 S 的元素：

$$
\begin{aligned}
&s_{11} \approx 1152.456, \quad s_{22} \approx 244.267, \quad s_{33} \approx 2915.556, \quad s_{44} \approx 197.067,\\
&s_{12} \approx -88.911, \quad s_{13} \approx 1589.667, \quad s_{14} \approx 301.600,\\
&s_{23} \approx 102.333, \quad s_{24} \approx -101.756, \quad s_{34} \approx 233.667.
\end{aligned}
$$

样本相关系数 $r_{jk} = \dfrac{s_{jk}}{\sqrt{s_{jj}s_{kk}}}$ 计算得：

$$
\begin{aligned}
&r_{12} \approx -0.1676, \quad r_{13} \approx 0.8673, \quad r_{14} \approx 0.6328,\\
&r_{23} \approx 0.1213, \quad r_{24} \approx -0.4638, \quad r_{34} \approx 0.3083.
\end{aligned}
$$

故样本相关矩阵为

$$
\boldsymbol{R} = \begin{pmatrix} 1 & -0.1676 & 0.8673 & 0.6328 \\ -0.1676 & 1 & 0.1213 & -0.4638 \\ 0.8673 & 0.1213 & 1 & 0.3083 \\ 0.6328 & -0.4638 & 0.3083 & 1 \end{pmatrix}.
$$

用R语言计算

```
# 创建数据矩阵
X <- matrix(c(
  230, 125, 200, 109,
  181, 99,  55,  107,
  165, 97,  105, 98,
  150, 115, 85,  71,
  97,  120, 0,   82,
  192, 100, 150, 103,
  181, 80,  85,  111,
  189, 90,  120, 93,
  172, 95,  110, 86,
  170, 125, 130, 78
), nrow = 10, byrow = TRUE)

colnames(X) <- c("销量", "价格变化", "广告费用", "促销员时长")

# 样本相关矩阵
R <- cor(X)
print(round(R, 4))
```

得到的相关矩阵为

	销量	价格变化	广告费用	促销员时长
销量	1.0000	−0.1676	0.8672	0.6329
价格变化	−0.1676	1.0000	0.1213	−0.4638
广告费用	0.8672	0.1213	1.0000	0.3083
促销员时长	0.6329	−0.4638	0.3083	1.0000

(b) [2 分] 就相关系数的符号进行说明.

【解】 根据计算所得相关系数：

- $r_{12} \approx -0.168$：负相关，表明销量与价格变化之间呈微弱负相关，即价格上升时销量略有下降。
- $r_{13} \approx 0.867$：强正相关，销量与广告费用高度正相关，广告投入增加可显著促进销售。
- $r_{14} \approx 0.633$：中等正相关，销量与促销员时长正相关，促销员投入有助于提升销量。
- $r_{23} \approx 0.121$：微弱正相关，价格变化与广告费用几乎无关。
- $r_{24} \approx -0.464$：中等负相关，价格变化与促销员时长负相关，可能促销员在价格下调时更有效。
- $r_{34} \approx 0.308$：弱正相关，广告费用与促销员时长有一定协同作用。

(c) [2 分] 检验假设 $\rho_{X_1 X_2} = 0$.

【解】 采用 t 检验，统计量

$$t = \frac{r_{12}\sqrt{n-2}}{\sqrt{1-r_{12}^2}} = \frac{-0.1676 \times \sqrt{8}}{\sqrt{1-0.0281}} \approx \frac{-0.1676 \times 2.828}{0.9858} \approx -0.481.$$

自由度 $df = n-2 = 8$，双侧检验 p-值 $\approx 2 \times P(T_8 > 0.481) \approx 2 \times 0.322 = 0.644$。由于 p-值 > 0.05，不拒绝原假设，即没有充分证据表明销量与价格变化之间存在线性相关关系。

6. [2 分] 证明 $\mathrm{rank}(\mathcal{H}) = \mathrm{tr}(\mathcal{H}) = n-1$, 其中 $\mathcal{H} = \mathcal{I}_p - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\mathrm{T}}$.

【证明】

$$\mathcal{H} = I_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^{\top}, \quad \mathbf{1}_n = (1,\ldots,1)^{\top}.$$

迹：

$$\mathrm{tr}(\mathcal{H}) = \mathrm{tr}(I_n) - \frac{1}{n}\mathrm{tr}(\mathbf{1}_n\mathbf{1}_n^{\top}) = n - \frac{1}{n}\cdot n = n-1.$$

幂等性：

$$\mathcal{H}^2 = I_n - \frac{2}{n}\mathbf{1}_n\mathbf{1}_n^{\top} + \frac{1}{n^2}(\mathbf{1}_n\mathbf{1}_n^{\top})^2 = I_n - \frac{2}{n}J_n + \frac{1}{n^2}nJ_n = I_n - \frac{1}{n}J_n = \mathcal{H},$$

其中 $J_n = \mathbf{1}_n\mathbf{1}_n^{\top}$ 且 $J_n^2 = nJ_n$。

秩： 幂等矩阵的秩等于其迹，故

$$\mathrm{rank}(\mathcal{H}) = \mathrm{tr}(\mathcal{H}) = n - 1.$$

7. 设 $\mathcal{X}$ 表示课堂中讨论过的钞票数据集当中伪钞数据的观测矩阵.

(a) [2 分] 计算 $\mathcal{X}$ 的样本协方差矩阵 $\mathcal{S} = \mathrm{Cov}(\mathcal{X})$.

【解】 代码如下

```
library(mclust)
data(banknote)
# 提取伪钞数据（变量：Length，Left，Right，Bottom，Top，Diagonal）
counterfeit <- banknote[banknote$Status == "counterfeit", 1:6]  # 前6列为测量变量
# 样本协方差矩阵
S <- cov(counterfeit)
print(round(S, 4))
```

运行得到：

样本协方差矩阵 $\mathcal{S} =$

	Length	Left	Right	Bottom	Top	Diagonal
Length	0.1418	0.0314	0.0231	−0.1032	−0.0185	0.0843
Left	0.0314	0.1303	0.1084	0.2158	0.1050	−0.2093
Right	0.0231	0.1084	0.1633	0.2841	0.1300	−0.2405
Bottom	−0.1032	0.2158	0.2841	2.0869	0.1645	−1.0370
Top	−0.0185	0.1050	0.1300	0.1645	0.6447	−0.5496
Diagonal	0.0843	−0.2093	−0.2405	−1.0370	−0.5496	1.3277

(b) [2 分] 作 $\mathcal{S}$ 的 Jordan 分解.

【解】 由于 $\mathcal{S}$ 是对称矩阵，其 Jordan 分解即为谱分解（正交对角化）。代码如下

```
   eig <- eigen(S)
round(eig$values, 4)
round(eig$vectors, 4)
```

根据 R 计算结果，特征值为：

$$\lambda_1 = 3.0003,\ \lambda_2 = 0.9356,\ \lambda_3 = 0.2434,\ \lambda_4 = 0.1947,\ \lambda_5 = 0.0852,\ \lambda_6 = 0.0355.$$

特征向量矩阵（列对应 $\lambda_1, \ldots, \lambda_6$）为：

$$\boldsymbol{V} \approx \begin{pmatrix} -0.0438 & 0.0107 & -0.3263 & 0.5617 & 0.7526 & 0.0981 \\ 0.1122 & 0.0714 & -0.2590 & 0.4555 & -0.3468 & -0.7665 \\ 0.1392 & 0.0663 & -0.3447 & 0.4153 & -0.5347 & 0.6317 \\ 0.7683 & -0.5631 & -0.2180 & -0.1861 & 0.1000 & -0.0222 \\ 0.2018 & 0.6593 & -0.5567 & -0.4507 & 0.1019 & -0.0349 \\ -0.5789 & -0.4885 & -0.5918 & -0.2584 & -0.0845 & -0.0457 \end{pmatrix}.$$

因此，$\mathcal{S}$ 的 Jordan 分解为 $\mathcal{S} = \boldsymbol{V}\,\mathrm{diag}(3.0003,\ 0.9356,\ 0.2434,\ 0.1947,\ 0.0852,\ 0.0355)\boldsymbol{V}^\top$。

(c) [2 分] 为什么所有的特征值均为正?

【解】 样本协方差矩阵定义为

$$\mathcal{S} = \frac{1}{n-1}\sum_{i=1}^{n}(\boldsymbol{x}_i - \bar{\boldsymbol{x}})(\boldsymbol{x}_i - \bar{\boldsymbol{x}})^\top = \frac{1}{n-1}\boldsymbol{X}_c^\top \boldsymbol{X}_c,$$

其中 $\boldsymbol{X}_c$ 是 $n \times p$ 的中心化数据矩阵。对于任意非零向量 $\boldsymbol{u} \in \mathbb{R}^p$，

$$\boldsymbol{u}^\top \mathcal{S} \boldsymbol{u} = \frac{1}{n-1}\|\boldsymbol{X}_c \boldsymbol{u}\|^2 \geq 0,$$

故 $\mathcal{S}$ 半正定。若 $\boldsymbol{X}_c$ 列满秩（即 $\mathrm{rank}(\boldsymbol{X}_c) = p$），则 $\boldsymbol{X}_c\boldsymbol{u} \neq \boldsymbol{0}$ 对所有 $\boldsymbol{u} \neq \boldsymbol{0}$ 成立，从而 $\boldsymbol{u}^\top \mathcal{S} \boldsymbol{u} > 0$，即 $\mathcal{S}$ 正定，所有特征值均为正。

在伪钞数据中，$n = 100$，$p = 6$，且六个测量变量（长度、左宽、右宽、底部、顶部、对角线）之间不存在精确线性关系，因此 $\boldsymbol{X}_c$ 列满秩，$\mathcal{S}$ 正定。计算结果中所有特征值均大于 0，验证了这一结论。