Statistics and Application 统计学与应用**, 2017, 6(1), 34-62**
Published Online March 2017 in Hans. http://www.hanspub.org/journal/sa
https://doi.org/10.12677/sa.2017.61005

Hans汉斯

The Sample Covariance Matrix and the Sample Correlation Matrix and Their Applications in the Sample Principal Component

Yingying Zhang, Tengzhong Rong

Department of Statistics and Actuarial Science, College of Mathematics and Statistics, Chongqing University, Chongqing
Email: robertzhang@cqu.edu.cn, timepdf@163.com

Received: Feb. 25th, 2017; accepted: Mar. 14th, 2017; published: Mar. 17th, 2017

Abstract

We give the properties and proofs of the sample principal component, and discuss them in two different conditions: from S on to calculate principal component and from R on to calculate principal component. From S on to calculate principal component, we give 7 properties (S1)-(S7) and their proofs, and the relationships stated by these properties get full display in Figure 1. Similarly, from R on to calculate principal component, we give 7 properties (R1)-(R7) and their proofs, and the relationships stated by these properties get full display in Figure 2. Finally we give two numerical simulation examples to verify the correctness of properties (S1)-(S7) and (R1)-(R7).

Keywords

Sample Covariance Matrix, Sample Correlation Matrix, Sample Principal Component, Properties and Proofs, R Software

样本协方差矩阵和样本相关矩阵及其在样本主成分中的应用

张应应，荣腾中

重庆大学数学与统计学院统计与精算学系，重庆
Email: robertzhang@cqu.edu.cn, timepdf@163.com

收稿日期：2017年2月25日；录用日期：2017年3月14日；发布日期：2017年3月17日

文章引用：张应应，荣腾中．样本协方差矩阵和样本相关矩阵及其在样本主成分中的应用[J]．统计学与应用，2017, 6(1): 34-62. https://doi.org/10.12677/sa.2017.61005

摘　要

我们给出了样本主成分的性质及证明，分两种情况讨论：从S出发求主成分和从R出发求主成分。在从S出发求主成分中，给出了7个性质(S1)-(S7)及它们的证明，这些性质说明的关系在图1中得到了充分的展现。同样，在从R出发求主成分中，给出了7个性质(R1)-(R7)及它们的证明，这些性质说明的关系在图2中得到了充分的展现。最后我们给出了两个数值模拟的例子来验证性质(S1)-(S7)和(R1)-(R7)的正确性。

关键词

样本协方差矩阵，样本相关矩阵，样本主成分，性质及证明，R软件

1. 引言

主成分分析(Principal Component Analysis, PCA)或称主分量分析，是一种降维的统计方法。通过正交变换将一组可能存在相关性的变量转换为一组线性不相关的综合变量，同时根据实际需要从中可以取出几个较少的综合变量尽可能多地反映原来变量的信息，转换后的这组综合变量叫主成分。主成分分析首先是由卡尔·皮尔森(Karl Pearson)对非随机变量引入的，尔后霍特林将此方法推广到随机向量的情形。信息的大小通常用方差来衡量。

为研究总体主成分和样本主成分的性质，需要定义总体或样本的协方差(矩阵)和相关系数(矩阵)。它们的定义及性质请参见[1]。总体主成分的性质在[2]-[10]中已有很好的总结及证明，但是他们对样本主成分的性质讨论却不多。由于总体的协方差矩阵和相关系数矩阵一般是未知的，而样本资料阵一般是已知的，人们一般采用样本主成分进行计算，从而研究样本主成分的性质就显得非常重要。本文给出了样本主成分的性质及证明，分两种情况讨论：从 $\boldsymbol{S}$ 出发求主成分并给出了 7 个性质(S1)-(S7)及它们的证明；从 $\boldsymbol{R}$ 出发求主成分并给出了 7 个性质(R1)-(R7)及它们的证明。

本文剩余部分安排如下：第 2 节给出样本主成分的性质及证明，分两种情况讨论：从 $\boldsymbol{S}$ 出发求主成分和从 $\boldsymbol{R}$ 出发求主成分。这些性质(S1)-(S7)和(R1)-(R7)说明的关系在图 1 和图 2 中得到了充分的展现。第 3 节给出了两个数值模拟的例子来验证性质(S1)-(S7)和(R1)-(R7)的正确性。第 4 节总结。

2. 样本主成分的性质及证明

有总体主成分和样本主成分，有协方差矩阵和相关矩阵，把这两个概念结合在一起形成表 1。本文讨论样本主成分的性质及证明。分两种情况讨论：从 $\boldsymbol{S}$ 或 $\boldsymbol{R}$ 出发求主成分。

2.1. 从 S 出发求主成分

设 $\hat{\lambda}_1 \ge \hat{\lambda}_2 \ge \cdots \ge \hat{\lambda}_p \ge 0$ 为样本协方差矩阵 $\boldsymbol{S}$ 的特征值，$\hat{\boldsymbol{a}}_1, \hat{\boldsymbol{a}}_2, \cdots, \hat{\boldsymbol{a}}_p$ 为相应的两两正交的单位特征向量，即

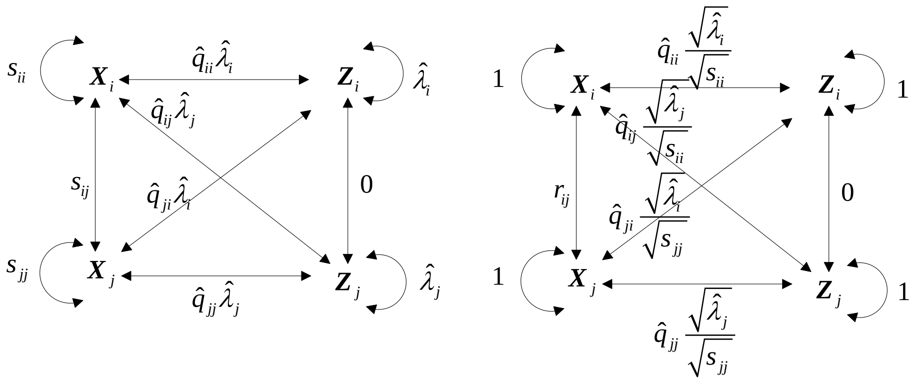

Figure 1. The relationships of the sample covariances (left) and the relationships of the sample correlations (right) among X_i, X_j, Z_i, Z_j

图 1. X_i, X_j, Z_i, Z_j 的样本协方差关系图(左)和样本相关系数关系图(右)

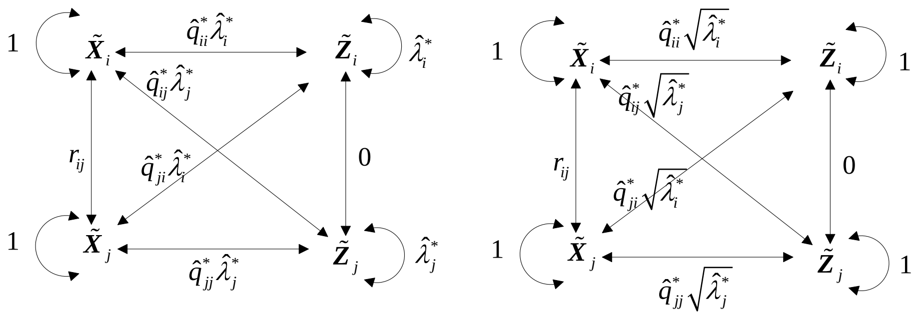

Figure 2. The relationships of the sample covariances (left) and the relationships of the sample correlations (right) among $\tilde{X}_i, \tilde{X}_j, \tilde{Z}_i, \tilde{Z}_j$

图 2. $\tilde{X}_i, \tilde{X}_j, \tilde{Z}_i, \tilde{Z}_j$ 的样本协方差关系图(左)和样本相关系数关系图(右)

Table 1. The covariance matrix and the correlation matrix of the population and the sample
表 1. 总体和样本的协方差矩阵和相关矩阵

	协方差矩阵	相关矩阵
总体	$\boldsymbol{\Sigma} = (\sigma_{ij})_{p\times p}$	$\boldsymbol{\rho} = (\rho_{ij})_{p\times p}$
样本	$\boldsymbol{S} = (s_{ij})_{p\times p}$	$\boldsymbol{R} = (r_{ij})_{p\times p}$

$$\boldsymbol{S}\hat{\boldsymbol{a}}_j = \hat{\lambda}_j \hat{\boldsymbol{a}}_j,$$
$$\hat{\boldsymbol{a}}_j^{\mathrm{T}}\hat{\boldsymbol{a}}_j = 1,\quad j = 1,2,\cdots,p,$$
$$\hat{\boldsymbol{a}}_i^{\mathrm{T}}\hat{\boldsymbol{a}}_j = 0,\quad i \ne j,\quad i,j = 1,2,\cdots,p.$$

令 $\hat{\boldsymbol{Q}} = (\hat{\boldsymbol{a}}_1, \hat{\boldsymbol{a}}_2, \cdots, \hat{\boldsymbol{a}}_p) = (\hat{q}_{ij})_{p\times p}$，它是一个正交阵，$\hat{\boldsymbol{\Lambda}} = \mathrm{diag}(\hat{\lambda}_1, \hat{\lambda}_2, \cdots, \hat{\lambda}_p)$，写成矩阵形式，就是

$$\boldsymbol{S}\hat{\boldsymbol{Q}} = \hat{\boldsymbol{Q}}\hat{\boldsymbol{\Lambda}},$$

上式两边左乘以 $\hat{\boldsymbol{Q}}^{\mathrm{T}}$，得到

$$\hat{\boldsymbol{Q}}^{\mathrm{T}}\boldsymbol{S}\hat{\boldsymbol{Q}} = \hat{\boldsymbol{\Lambda}}.$$

则第 i 个主成分 $z_i = \hat{\boldsymbol{a}}_i^{\mathrm{T}}\boldsymbol{x}$，$i = 1,2,\cdots,p$，其中 $\boldsymbol{x} = (x_1, x_2, \cdots, x_p)^{\mathrm{T}}$，且

$$z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_p \end{pmatrix} = \begin{pmatrix} \hat{\boldsymbol{a}}_1^{\mathrm{T}} \boldsymbol{x} \\ \hat{\boldsymbol{a}}_2^{\mathrm{T}} \boldsymbol{x} \\ \vdots \\ \hat{\boldsymbol{a}}_p^{\mathrm{T}} \boldsymbol{x} \end{pmatrix} = \begin{pmatrix} \hat{\boldsymbol{a}}_1^{\mathrm{T}} \\ \hat{\boldsymbol{a}}_2^{\mathrm{T}} \\ \vdots \\ \hat{\boldsymbol{a}}_p^{\mathrm{T}} \end{pmatrix} \boldsymbol{x} = \left(\hat{\boldsymbol{a}}_1, \hat{\boldsymbol{a}}_2, \cdots, \hat{\boldsymbol{a}}_p \right)^{\mathrm{T}} \boldsymbol{x} = \hat{\boldsymbol{Q}}^{\mathrm{T}} \boldsymbol{x}.$$

下面构造样本主成分，令

$$\boldsymbol{Z}_{(k)} = \hat{\boldsymbol{Q}}^{\mathrm{T}} \boldsymbol{X}_{(k)}.$$

因此样本主成分为

$$\boldsymbol{Z} = \begin{pmatrix} z_{11} & z_{12} & \cdots & z_{1p} \\ z_{21} & z_{22} & \cdots & z_{2p} \\ \vdots & \vdots & & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{np} \end{pmatrix} = \begin{pmatrix} \boldsymbol{Z}_{(1)}^{\mathrm{T}} \\ \boldsymbol{Z}_{(2)}^{\mathrm{T}} \\ \vdots \\ \boldsymbol{Z}_{(n)}^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} \boldsymbol{X}_{(1)}^{\mathrm{T}} \hat{\boldsymbol{Q}} \\ \boldsymbol{X}_{(2)}^{\mathrm{T}} \hat{\boldsymbol{Q}} \\ \vdots \\ \boldsymbol{X}_{(n)}^{\mathrm{T}} \hat{\boldsymbol{Q}} \end{pmatrix} = \boldsymbol{X} \hat{\boldsymbol{Q}}$$

$$= \left(\boldsymbol{X} \hat{\boldsymbol{a}}_1, \boldsymbol{X} \hat{\boldsymbol{a}}_2, \cdots, \boldsymbol{X} \hat{\boldsymbol{a}}_p \right) = \left(\boldsymbol{Z}_1, \boldsymbol{Z}_2, \cdots, \boldsymbol{Z}_p \right),$$

其中 $\boldsymbol{Z}_{(k)}^{\mathrm{T}} = \left(z_{k1}, z_{k2}, \cdots, z_{kp} \right)$ 表示样本主成分的各行， $\boldsymbol{Z}_j = \left(z_{1j}, z_{2j}, \cdots, z_{nj} \right)^{\mathrm{T}}$ 表示样本主成分的各列。易知

$$\bar{\boldsymbol{Z}} = \frac{1}{n} \sum_{k=1}^{n} \boldsymbol{Z}_{(k)} = \frac{1}{n} \sum_{k=1}^{n} \hat{\boldsymbol{Q}}^{\mathrm{T}} \boldsymbol{X}_{(k)} = \hat{\boldsymbol{Q}}^{\mathrm{T}} \frac{1}{n} \sum_{k=1}^{n} \boldsymbol{X}_{(k)} = \hat{\boldsymbol{Q}}^{\mathrm{T}} \bar{\boldsymbol{X}}.$$

类似于数据资料阵 $\boldsymbol{X}$ 的标准化[1]，可得到矩阵 $\boldsymbol{Z}$ 的标准化矩阵 $\boldsymbol{Z}^*$ 满足

$$\boldsymbol{Z}^* = \left(z_{ij}^* \right)_{n \times p} = \begin{pmatrix} \boldsymbol{Z}_{(1)}^{*T} \\ \boldsymbol{Z}_{(2)}^{*T} \\ \vdots \\ \boldsymbol{Z}_{(n)}^{*T} \end{pmatrix} = \left(\boldsymbol{Z}_1^*, \boldsymbol{Z}_2^*, \cdots, \boldsymbol{Z}_p^* \right),$$

其中

$$\boldsymbol{Z}_{(k)}^* = \left(\frac{z_{k1} - \bar{z}_1}{\sqrt{\hat{\lambda}_1}}, \frac{z_{k2} - \bar{z}_2}{\sqrt{\hat{\lambda}_2}}, \cdots, \frac{z_{kp} - \bar{z}_p}{\sqrt{\hat{\lambda}_p}} \right)^{\mathrm{T}}$$

$$= \begin{pmatrix} \frac{1}{\sqrt{\hat{\lambda}_1}} & & & \\ & \frac{1}{\sqrt{\hat{\lambda}_2}} & & \\ & & \ddots & \\ & & & \frac{1}{\sqrt{\hat{\lambda}_p}} \end{pmatrix} \begin{pmatrix} z_{k1} - \bar{z}_1 \\ z_{k2} - \bar{z}_2 \\ \vdots \\ z_{kp} - \bar{z}_p \end{pmatrix}$$

$$= \hat{\boldsymbol{\Lambda}}^{-\frac{1}{2}} \left(\boldsymbol{Z}_{(k)} - \bar{\boldsymbol{Z}} \right),$$

$$\hat{\boldsymbol{\Lambda}} = \mathrm{diag}\left(\hat{\lambda}_1, \hat{\lambda}_2, \cdots, \hat{\lambda}_p \right) = \mathrm{cov}\left(\boldsymbol{Z} \right),$$

上式中 $\mathrm{diag}\left(\hat{\lambda}_1, \hat{\lambda}_2, \cdots, \hat{\lambda}_p \right)$ 表示由向量 $\left(\hat{\lambda}_1, \hat{\lambda}_2, \cdots, \hat{\lambda}_p \right)^{\mathrm{T}}$ 作为对角线元素构成的对角阵。注意，上式用到了(S2)

的结论。从而

$$\boldsymbol{Z}^{*}=\begin{pmatrix}\boldsymbol{Z}_{(1)}^{*\mathrm{T}}\\ \boldsymbol{Z}_{(2)}^{*\mathrm{T}}\\ \vdots\\ \boldsymbol{Z}_{(n)}^{*\mathrm{T}}\end{pmatrix}=\begin{pmatrix}\left(\boldsymbol{Z}_{(1)}-\overline{\boldsymbol{Z}}\right)^{\mathrm{T}}\hat{\Lambda}^{-\frac{1}{2}}\\ \left(\boldsymbol{Z}_{(2)}-\overline{\boldsymbol{Z}}\right)^{\mathrm{T}}\hat{\Lambda}^{-\frac{1}{2}}\\ \vdots\\ \left(\boldsymbol{Z}_{(n)}-\overline{\boldsymbol{Z}}\right)^{\mathrm{T}}\hat{\Lambda}^{-\frac{1}{2}}\end{pmatrix}=\begin{pmatrix}\boldsymbol{Z}_{(1)}^{\mathrm{T}}-\overline{\boldsymbol{Z}}^{\mathrm{T}}\\ \boldsymbol{Z}_{(2)}^{\mathrm{T}}-\overline{\boldsymbol{Z}}^{\mathrm{T}}\\ \vdots\\ \boldsymbol{Z}_{(n)}^{\mathrm{T}}-\overline{\boldsymbol{Z}}^{\mathrm{T}}\end{pmatrix}\hat{\Lambda}^{-\frac{1}{2}}=\left(\boldsymbol{Z}-\boldsymbol{1}_{n\times 1}\overline{\boldsymbol{Z}}^{\mathrm{T}}\right)\hat{\Lambda}^{-\frac{1}{2}}.$$

易知

$$\overline{\boldsymbol{Z}^{*}}=\frac{1}{n}\sum_{k=1}^{n}\boldsymbol{Z}_{(k)}^{*}=\frac{1}{n}\sum_{k=1}^{n}\hat{\Lambda}^{-\frac{1}{2}}\left(\boldsymbol{Z}_{(k)}-\overline{\boldsymbol{Z}}\right)=\frac{1}{n}\hat{\Lambda}^{-\frac{1}{2}}\left(n\overline{\boldsymbol{Z}}-n\overline{\boldsymbol{Z}}\right)=\boldsymbol{0}.$$

因为 $\boldsymbol{Z}_{(k)}=\hat{\boldsymbol{Q}}^{\mathrm{T}}\boldsymbol{X}_{(k)}$，人们自然会想到 $\boldsymbol{Z}_{(k)}^{*}=\hat{\boldsymbol{Q}}^{\mathrm{T}}\boldsymbol{X}_{(k)}^{*}$，但是一般来说这是错误的。详细的证明请见本文的附录。

我们用 $\boldsymbol{S}\geq\boldsymbol{0}$ 表示它是一个非负定阵，用 $\boldsymbol{S}>\boldsymbol{0}$ 表示它是一个正定阵。令 $\boldsymbol{X}=\left(x_{ij}\right)_{n\times p}$ 为数据资料矩阵，$\boldsymbol{Z}=\left(z_{ij}\right)_{n\times p}$ 为主成分得分矩阵，则对于由 $\boldsymbol{S}\geq\boldsymbol{0}$ 出发计算的样本主成分有如下性质：

(S1).

$$\operatorname{cov}(\boldsymbol{X})=\operatorname{cov}(\boldsymbol{X},\boldsymbol{X})=\left(\operatorname{cov}\left(\boldsymbol{X}_{i},\boldsymbol{X}_{j}\right)\right)_{p\times p}=\boldsymbol{S}=\left(s_{ij}\right)_{p\times p},$$

特别地，

$$\operatorname{cov}\left(\boldsymbol{X}_{i},\boldsymbol{X}_{j}\right)=s_{ij},\quad i,j=1,\cdots,p.$$

(S2).

$$\operatorname{cov}(\boldsymbol{Z})=\operatorname{cov}(\boldsymbol{Z},\boldsymbol{Z})=\left(\operatorname{cov}\left(\boldsymbol{Z}_{i},\boldsymbol{Z}_{j}\right)\right)_{p\times p}=\hat{\Lambda},$$

特别地，

$$\operatorname{cov}\left(\boldsymbol{Z}_{j}\right)=\operatorname{cov}\left(\boldsymbol{Z}_{j},\boldsymbol{Z}_{j}\right)=\hat{\lambda}_{j},\quad j=1,\cdots,p,$$
$$\operatorname{cov}\left(\boldsymbol{Z}_{i},\boldsymbol{Z}_{j}\right)=0,\quad i\neq j,\quad i,j=1,\cdots,p.$$

(S3).

$$\operatorname{cov}(\boldsymbol{X},\boldsymbol{Z})=\left(\operatorname{cov}\left(\boldsymbol{X}_{i},\boldsymbol{Z}_{j}\right)\right)_{p\times p}=\left(\hat{q}_{ij}\hat{\lambda}_{j}\right)_{p\times p}=\hat{\boldsymbol{Q}}\hat{\Lambda},$$

特别地，

$$\operatorname{cov}\left(\boldsymbol{X}_{i},\boldsymbol{Z}_{j}\right)=\hat{q}_{ij}\hat{\lambda}_{j},\quad i,j=1,\cdots,p.$$

(S4).

$$\begin{aligned}\operatorname{cor}(\boldsymbol{X})&=\operatorname{cor}(\boldsymbol{X},\boldsymbol{X})=\left(\operatorname{cor}\left(\boldsymbol{X}_{i},\boldsymbol{X}_{j}\right)\right)_{p\times p}\\&=\operatorname{cov}\left(\boldsymbol{X}^{*}\right)=\left(\operatorname{cov}\left(\boldsymbol{X}_{i}^{*},\boldsymbol{X}_{j}^{*}\right)\right)_{p\times p}=\boldsymbol{R}=\left(r_{ij}\right)_{p\times p},\end{aligned}$$

特别地，

$$\operatorname{cor}\left(\boldsymbol{X}_{i},\boldsymbol{X}_{j}\right)=\operatorname{cov}\left(\boldsymbol{X}_{i}^{*},\boldsymbol{X}_{j}^{*}\right)=r_{ij},\quad i,j=1,\cdots,p.$$

(S5). 若 $\boldsymbol{S}>\boldsymbol{0}$，则

$$\begin{aligned}\operatorname{cor}(\boldsymbol{Z})&=\operatorname{cor}(\boldsymbol{Z},\boldsymbol{Z})=\left(\operatorname{cor}\left(\boldsymbol{Z}_i,\boldsymbol{Z}_j\right)\right)_{p\times p}=\operatorname{cov}\left(Z^*\right)=\left(\operatorname{cov}\left(\boldsymbol{Z}_i^*,\boldsymbol{Z}_j^*\right)\right)_{p\times p}\\&=\boldsymbol{E}_{p\times p}=\operatorname{diag}(1,1,\cdots,1)_{p\times p},\end{aligned}$$

特别地，

$$\operatorname{cor}\left(\boldsymbol{Z}_j\right)=\operatorname{cov}\left(\boldsymbol{Z}_j^*\right)=\operatorname{cov}\left(\boldsymbol{Z}_j^*,\boldsymbol{Z}_j^*\right)=1,\quad j=1,\cdots,p,$$
$$\operatorname{cor}\left(\boldsymbol{Z}_i,\boldsymbol{Z}_j\right)=\operatorname{cov}\left(\boldsymbol{Z}_i^*,\boldsymbol{Z}_j^*\right)=0,\quad i\neq j,\quad i,j=1,\cdots,p.$$

(S6). 若 $\boldsymbol{S}>\boldsymbol{0}$，则

$$\begin{aligned}\operatorname{cor}(\boldsymbol{X},\boldsymbol{Z})&=\left(\operatorname{cor}\left(\boldsymbol{X}_i,\boldsymbol{Z}_j\right)\right)_{p\times p}=\operatorname{cov}\left(\boldsymbol{X}^*,\boldsymbol{Z}^*\right)=\left(\operatorname{cov}\left(\boldsymbol{X}_i^*,\boldsymbol{Z}_j^*\right)\right)_{p\times p}\\&=\left(\hat{q}_{ij}\frac{\sqrt{\hat{\lambda}_j}}{\sqrt{s_{ii}}}\right)_{p\times p}=\hat{\boldsymbol{D}}^{-\frac{1}{2}}\hat{\boldsymbol{Q}}\hat{\boldsymbol{\Lambda}}^{\frac{1}{2}},\end{aligned}$$

其中

$$\hat{\boldsymbol{D}}=\operatorname{diag}\left(s_{11},s_{22},\cdots,s_{pp}\right)$$

是 $\boldsymbol{S}$ 的对角线元素组成的对角阵。特别地，

$$\operatorname{cor}\left(\boldsymbol{X}_i,\boldsymbol{Z}_j\right)=\operatorname{cov}\left(\boldsymbol{X}_i^*,\boldsymbol{Z}_j^*\right)=\hat{q}_{ij}\frac{\sqrt{\hat{\lambda}_j}}{\sqrt{s_{ii}}},\quad i,j=1,\cdots,p.$$

(S7). 样本总方差

$$\sum_{j=1}^{p}\operatorname{cov}\left(\boldsymbol{X}_j\right)=\sum_{j=1}^{p}s_{jj}=\sum_{j=1}^{p}\hat{\lambda}_j=\sum_{j=1}^{p}\operatorname{cov}\left(\boldsymbol{Z}_j\right).$$

此式表明了样本 $\boldsymbol{X}_1,\cdots,\boldsymbol{X}_p$ 的样本(协)方差之和等于样本主成分 $\boldsymbol{Z}_1,\cdots,\boldsymbol{Z}_p$ 的样本(协)方差之和。

在以上 7 个性质中，(S1)显然，(S2)和(S6)的分量形式的结果和(S7)可以在[7]中找到。

注意，在(S5)和(S6)中我们要求 $\boldsymbol{S}>\boldsymbol{0}$，而其余性质只要求 $\boldsymbol{S}\geq\boldsymbol{0}$。因为在(S5)和(S6)的证明中会涉及到

$$\hat{\boldsymbol{\Lambda}}^{-\frac{1}{2}}=\operatorname{diag}\left(\frac{1}{\sqrt{\hat{\lambda}_1}},\frac{1}{\sqrt{\hat{\lambda}_2}},\cdots,\frac{1}{\sqrt{\hat{\lambda}_p}}\right),$$

若 $\boldsymbol{S}$ 有一个 0 特征值，即 $\hat{\lambda}_p=0$，则 $1/\sqrt{\hat{\lambda}_p}$ 是没有意义的，从而结论不成立，数值结果也证实了这一点。

由性质(S1)-(S6)归纳得到样本 $\boldsymbol{X}_i,\boldsymbol{X}_j,\boldsymbol{Z}_i,\boldsymbol{Z}_j$ 的样本协方差关系图((S1)-(S3))和样本相关系数关系图((S4)-(S6))，如图 1。

性质(S1)-(S7)的证明有些是初等的，有些需要一定的技巧，由于这些证明较占篇幅，所以把它们放到了本文的附录中。

2.2. 从 R 出发求主成分

设 $\hat{\lambda}_1^*\geq\hat{\lambda}_2^*\geq\cdots\geq\hat{\lambda}_p^*\geq0$ 为样本相关矩阵 $\boldsymbol{R}$ 的特征值，$\hat{\boldsymbol{a}}_1^*,\hat{\boldsymbol{a}}_2^*,\cdots,\hat{\boldsymbol{a}}_p^*$ 为相应的两两正交的单位特征向量，

即

$$R\hat{a}_j^* = \hat{\lambda}_j^* \hat{a}_j^*,$$
$$\hat{a}_j^{*T}\hat{a}_j^* = 1,\quad j = 1,2,\cdots,p,$$
$$\hat{a}_i^{*T}\hat{a}_j^* = 0,\quad i \neq j,\quad i,j = 1,2,\cdots,p.$$

令 $\hat{Q}^* = \left(\hat{a}_1^*, \hat{a}_2^*, \cdots, \hat{a}_p^*\right) = \left(\hat{q}_{ij}^*\right)_{p\times p}$，它是一个正交阵，$\hat{\Lambda}^* = \mathrm{diag}\left(\hat{\lambda}_1^*, \hat{\lambda}_2^*, \cdots, \hat{\lambda}_p^*\right)$，写成矩阵形式，就是

$$R\hat{Q}^* = \hat{Q}^*\hat{\Lambda}^*,$$

上式两边左乘以 $\hat{Q}^{*\mathrm{T}}$，得到

$$\hat{Q}^{*\mathrm{T}} R \hat{Q}^* = \hat{\Lambda}^*.$$

则第 i 个主成分 $\tilde{z}_i = \hat{a}_i^{*\mathrm{T}}\tilde{x},\ i = 1,2,\cdots,p$，其中 $\tilde{x} = x^* = \left(x_1^*, x_2^*, \cdots, x_p^*\right)^{\mathrm{T}}$。这是因为我们从样本相关矩阵 R 出发，所以变量已经标准化了。且

$$\tilde{z} = \begin{pmatrix} \tilde{z}_1 \\ \tilde{z}_2 \\ \vdots \\ \tilde{z}_p \end{pmatrix} = \begin{pmatrix} \hat{a}_1^{*\mathrm{T}}\tilde{x} \\ \hat{a}_2^{*\mathrm{T}}\tilde{x} \\ \vdots \\ \hat{a}_p^{*\mathrm{T}}\tilde{x} \end{pmatrix} = \begin{pmatrix} \hat{a}_1^{*\mathrm{T}} \\ \hat{a}_2^{*\mathrm{T}} \\ \vdots \\ \hat{a}_p^{*\mathrm{T}} \end{pmatrix}\tilde{x} = \left(\hat{a}_1^*, \hat{a}_2^*, \cdots, \hat{a}_p^*\right)^{\mathrm{T}}\tilde{x} = \hat{Q}^{*\mathrm{T}}\tilde{x} = \hat{Q}^{*\mathrm{T}}x^*.$$

下面构造样本主成分，令

$$\tilde{Z}_{(k)} = \hat{Q}^{*\mathrm{T}}\tilde{X}_{(k)} = \hat{Q}^{*\mathrm{T}}X_{(k)}^*,$$

则

$$\tilde{X}_{(k)} = \hat{Q}^*\tilde{Z}_{(k)}.$$

因此样本主成分为

$$\tilde{Z} = \begin{pmatrix} \tilde{z}_{11} & \tilde{z}_{12} & \cdots & \tilde{z}_{1p} \\ \tilde{z}_{21} & \tilde{z}_{22} & \cdots & \tilde{z}_{2p} \\ \vdots & \vdots & & \vdots \\ \tilde{z}_{n1} & \tilde{z}_{n2} & \cdots & \tilde{z}_{np} \end{pmatrix} = \begin{pmatrix} \tilde{Z}_{(1)}^{\mathrm{T}} \\ \tilde{Z}_{(2)}^{\mathrm{T}} \\ \vdots \\ \tilde{Z}_{(n)}^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} \tilde{X}_{(1)}^{\mathrm{T}}\hat{Q}^* \\ \tilde{X}_{(2)}^{\mathrm{T}}\hat{Q}^* \\ \vdots \\ \tilde{X}_{(n)}^{\mathrm{T}}\hat{Q}^* \end{pmatrix} = \begin{pmatrix} \tilde{X}_{(1)}^{\mathrm{T}} \\ \tilde{X}_{(2)}^{\mathrm{T}} \\ \vdots \\ \tilde{X}_{(n)}^{\mathrm{T}} \end{pmatrix}\hat{Q}^*$$
$$= \tilde{X}\hat{Q}^* = X^*\hat{Q}^* = \left(\tilde{X}\hat{a}_1^*, \tilde{X}\hat{a}_2^*, \cdots, \tilde{X}\hat{a}_p^*\right) = \left(\tilde{Z}_1, \tilde{Z}_2, \cdots, \tilde{Z}_p\right),$$

其中 $\tilde{Z}_{(k)}^{\mathrm{T}} = \left(\tilde{z}_{k1}, \tilde{z}_{k2}, \cdots, \tilde{z}_{kp}\right)$ 表示样本主成分的各行，$\tilde{Z}_j = \left(\tilde{z}_{1j}, \tilde{z}_{2j}, \cdots, \tilde{z}_{nj}\right)^{\mathrm{T}}$ 表示样本主成分的各列。注意到

$$\tilde{X} = \begin{pmatrix} \tilde{X}_{(1)}^{\mathrm{T}} \\ \tilde{X}_{(2)}^{\mathrm{T}} \\ \vdots \\ \tilde{X}_{(n)}^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} X_{(1)}^{*\mathrm{T}} \\ X_{(2)}^{*\mathrm{T}} \\ \vdots \\ X_{(n)}^{*\mathrm{T}} \end{pmatrix} = X^*,$$

X^* 和 $X_{(k)}^*$ 可参照[1]中的数据资料矩阵 X 的标准化部分。但是，

$$\tilde{Z}_{(k)} = \hat{Q}^{*\mathrm{T}}X_{(k)}^* = \hat{Q}^{*\mathrm{T}}\hat{D}^{-\frac{1}{2}}\left(X_{(k)} - \bar{X}\right),$$
$$Z_{(k)}^* = \hat{\Lambda}^{-\frac{1}{2}}\left(Z_{(k)} - \bar{Z}\right) = \hat{\Lambda}^{-\frac{1}{2}}\hat{Q}^{\mathrm{T}}\left(X_{(k)} - \bar{X}\right),$$

所以 $\tilde{\boldsymbol{Z}}_{(k)} \neq \boldsymbol{Z}^*_{(k)}$，从而 $\tilde{\boldsymbol{Z}} \neq \boldsymbol{Z}^*$。易知

$$\overline{\tilde{\boldsymbol{X}}} = \overline{\boldsymbol{X}^*} = \boldsymbol{0}.$$

容易验证

$$\overline{\tilde{\boldsymbol{Z}}} = \frac{1}{n}\sum_{k=1}^{n}\tilde{\boldsymbol{Z}}_{(k)} = \frac{1}{n}\sum_{k=1}^{n}\hat{\boldsymbol{Q}}^{*\mathrm{T}}\tilde{\boldsymbol{X}}_{(k)} = \hat{\boldsymbol{Q}}^{*\mathrm{T}}\frac{1}{n}\sum_{k=1}^{n}\tilde{\boldsymbol{X}}_{(k)} = \hat{\boldsymbol{Q}}^{*\mathrm{T}}\overline{\tilde{\boldsymbol{X}}} = \boldsymbol{0}.$$

由 $\tilde{\boldsymbol{X}} = \boldsymbol{X}^*$ 和 $\overline{\boldsymbol{X}^*} = \boldsymbol{0}$ 有

$$\tilde{\boldsymbol{X}}^* = \boldsymbol{X}^{**} = \left(x_{ij}^{**}\right)_{n\times p} = \begin{pmatrix} \boldsymbol{X}_{(1)}^{**\mathrm{T}} \\ \boldsymbol{X}_{(2)}^{**\mathrm{T}} \\ \vdots \\ \boldsymbol{X}_{(n)}^{**\mathrm{T}} \end{pmatrix},$$

其中

$$\begin{aligned} \boldsymbol{X}_{(k)}^{**} &= \left(\frac{x_{k1}^* - \overline{x_1^*}}{\sqrt{r_{11}}}, \frac{x_{k2}^* - \overline{x_2^*}}{\sqrt{r_{22}}}, \cdots, \frac{x_{kp}^* - \overline{x_p^*}}{\sqrt{r_{pp}}}\right)^{\mathrm{T}} \\ &= \left(\frac{x_{k1}^* - 0}{1}, \frac{x_{k2}^* - 0}{1}, \cdots, \frac{x_{kp}^* - 0}{1}\right)^{\mathrm{T}} \\ &= \left(x_{k1}^*, x_{k2}^*, \cdots, x_{kp}^*\right)^{\mathrm{T}} = \boldsymbol{X}_{(k)}^*, \end{aligned}$$

注意，上式用到了(S4)的结论 $\operatorname{cov}\left(\boldsymbol{X}^*\right) = \boldsymbol{R}$。则

$$\tilde{\boldsymbol{X}}^* = \boldsymbol{X}^{**} = \boldsymbol{X}^* = \tilde{\boldsymbol{X}}.$$

易知

$$\begin{aligned} &\overline{\tilde{\boldsymbol{X}}^*} = \overline{\boldsymbol{X}^{**}} = \overline{\boldsymbol{X}^*} = \boldsymbol{0}, \\ &\tilde{\boldsymbol{X}}_{(k)}^* = \boldsymbol{X}_{(k)}^{**} = \boldsymbol{X}_{(k)}^* = \tilde{\boldsymbol{X}}_{(k)}, \quad k = 1, \cdots, n. \end{aligned}$$

类似于矩阵 $\boldsymbol{X}$ 的标准化，可得到矩阵 $\tilde{\boldsymbol{Z}}$ 的标准化矩阵 $\tilde{\boldsymbol{Z}}^*$ 满足

$$\tilde{\boldsymbol{Z}}^* = \left(\tilde{z}_{ij}^*\right)_{n\times p} = \begin{pmatrix} \tilde{\boldsymbol{Z}}_{(1)}^{*\mathrm{T}} \\ \tilde{\boldsymbol{Z}}_{(2)}^{*\mathrm{T}} \\ \vdots \\ \tilde{\boldsymbol{Z}}_{(n)}^{*\mathrm{T}} \end{pmatrix} = \left(\tilde{\boldsymbol{Z}}_1^*, \tilde{\boldsymbol{Z}}_2^*, \cdots, \tilde{\boldsymbol{Z}}_p^*\right),$$

其中

$$\begin{aligned} \tilde{\boldsymbol{Z}}_{(k)}^* &= \begin{pmatrix} \left(\tilde{z}_{k1} - \overline{\tilde{z}_1}\right)\Big/\sqrt{\hat{\lambda}_1^*} \\ \left(\tilde{z}_{k2} - \overline{\tilde{z}_2}\right)\Big/\sqrt{\hat{\lambda}_2^*} \\ \vdots \\ \left(\tilde{z}_{kp} - \overline{\tilde{z}_p}\right)\Big/\sqrt{\hat{\lambda}_p^*} \end{pmatrix} = \begin{pmatrix} 1\Big/\sqrt{\hat{\lambda}_1^*} & & & \\ & 1\Big/\sqrt{\hat{\lambda}_2^*} & & \\ & & \ddots & \\ & & & 1\Big/\sqrt{\hat{\lambda}_p^*} \end{pmatrix} \begin{pmatrix} \tilde{z}_{k1} - \overline{\tilde{z}_1} \\ \tilde{z}_{k2} - \overline{\tilde{z}_2} \\ \vdots \\ \tilde{z}_{kp} - \overline{\tilde{z}_p} \end{pmatrix} \\ &= \left(\hat{\boldsymbol{\Lambda}}^*\right)^{-\frac{1}{2}}\left(\tilde{\boldsymbol{Z}}_{(k)} - \overline{\tilde{\boldsymbol{Z}}}\right) = \left(\hat{\boldsymbol{\Lambda}}^*\right)^{-\frac{1}{2}}\tilde{\boldsymbol{Z}}_{(k)}, \\ \hat{\boldsymbol{\Lambda}}^* &= \operatorname{diag}\left(\hat{\lambda}_1^*, \hat{\lambda}_2^*, \cdots, \hat{\lambda}_p^*\right) = \operatorname{cov}\left(\tilde{\boldsymbol{Z}}\right). \end{aligned}$$

注意，上式用到了(R2)的结论。从而

$$\tilde{\boldsymbol{Z}}^{*}=\begin{pmatrix}\tilde{\boldsymbol{Z}}_{(1)}^{*\mathrm{T}}\\ \tilde{\boldsymbol{Z}}_{(2)}^{*\mathrm{T}}\\ \vdots\\ \tilde{\boldsymbol{Z}}_{(n)}^{*\mathrm{T}}\end{pmatrix}=\begin{pmatrix}\tilde{\boldsymbol{Z}}_{(1)}^{\mathrm{T}}\left(\hat{\Lambda}^{*}\right)^{-\frac{1}{2}}\\ \tilde{\boldsymbol{Z}}_{(2)}^{\mathrm{T}}\left(\hat{\Lambda}^{*}\right)^{-\frac{1}{2}}\\ \vdots\\ \tilde{\boldsymbol{Z}}_{(n)}^{\mathrm{T}}\left(\hat{\Lambda}^{*}\right)^{-\frac{1}{2}}\end{pmatrix}=\begin{pmatrix}\tilde{\boldsymbol{Z}}_{(1)}^{\mathrm{T}}\\ \tilde{\boldsymbol{Z}}_{(2)}^{\mathrm{T}}\\ \vdots\\ \tilde{\boldsymbol{Z}}_{(n)}^{\mathrm{T}}\end{pmatrix}\left(\hat{\Lambda}^{*}\right)^{-\frac{1}{2}}=\tilde{\boldsymbol{Z}}\left(\hat{\Lambda}^{*}\right)^{-\frac{1}{2}}.$$

易知

$$\begin{aligned}\overline{\tilde{\boldsymbol{Z}}^{*}}&=\frac{1}{n}\sum_{k=1}^{n}\tilde{\boldsymbol{Z}}_{(k)}^{*}=\frac{1}{n}\sum_{k=1}^{n}\left(\hat{\Lambda}^{*}\right)^{-\frac{1}{2}}\tilde{\boldsymbol{Z}}_{(k)}\\&=\left(\hat{\Lambda}^{*}\right)^{-\frac{1}{2}}\left(\frac{1}{n}\sum_{k=1}^{n}\tilde{\boldsymbol{Z}}_{(k)}\right)=\left(\hat{\Lambda}^{*}\right)^{-\frac{1}{2}}\overline{\tilde{\boldsymbol{Z}}}=\boldsymbol{0}.\end{aligned}$$

因为 $\tilde{\boldsymbol{Z}}_{(k)}=\hat{\boldsymbol{Q}}^{*\mathrm{T}}\tilde{\boldsymbol{X}}_{(k)}$，人们自然会想到 $\tilde{\boldsymbol{Z}}_{(k)}^{*}=\hat{\boldsymbol{Q}}^{*\mathrm{T}}\tilde{\boldsymbol{X}}_{(k)}^{*}$，同样，一般来说这是错误的。详细的证明请见本文的附录。

值得一提的是，矩阵 $\tilde{\boldsymbol{Z}}$ 和它的标准化矩阵 $\tilde{\boldsymbol{Z}}^{*}$ 的推导是新的。

我们用 $\boldsymbol{R}\geq\boldsymbol{0}$ 表示它是一个非负定阵，$\boldsymbol{R}>\boldsymbol{0}$ 表示它是一个正定阵。令 $\tilde{\boldsymbol{X}}=\left(\tilde{x}_{ij}\right)_{n\times p}=\boldsymbol{X}^{*}$ 为数据资料矩阵，$\tilde{\boldsymbol{Z}}=\left(\tilde{z}_{ij}\right)_{n\times p}$ 为主成分得分矩阵，则对于由 $\boldsymbol{R}\geq\boldsymbol{0}$ 出发计算的样本主成分有如下性质：

(R1).

$$\begin{aligned}\operatorname{cov}\left(\tilde{\boldsymbol{X}}\right)=\operatorname{cov}\left(\boldsymbol{X}^{*}\right)=\operatorname{cor}\left(\boldsymbol{X}\right)=\boldsymbol{R}=\left(r_{ij}\right)_{p\times p}\\=\operatorname{cov}\left(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{X}}\right)=\left(\operatorname{cov}\left(\tilde{\boldsymbol{X}}_{i},\tilde{\boldsymbol{X}}_{j}\right)\right)_{p\times p},\end{aligned}$$

特别地，

$$\operatorname{cov}\left(\tilde{\boldsymbol{X}}_{i},\tilde{\boldsymbol{X}}_{j}\right)=\operatorname{cov}\left(\boldsymbol{X}_{i}^{*},\boldsymbol{X}_{j}^{*}\right)=r_{ij},\quad i,j=1,\cdots,p.$$

(R2).

$$\operatorname{cov}\left(\tilde{\boldsymbol{Z}}\right)=\operatorname{cov}\left(\tilde{\boldsymbol{Z}},\tilde{\boldsymbol{Z}}\right)=\left(\operatorname{cov}\left(\tilde{\boldsymbol{Z}}_{i},\tilde{\boldsymbol{Z}}_{j}\right)\right)_{p\times p}=\hat{\Lambda}^{*},$$

特别地，

$$\begin{aligned}&\operatorname{cov}\left(\tilde{\boldsymbol{Z}}_{j}\right)=\operatorname{cov}\left(\tilde{\boldsymbol{Z}}_{j},\tilde{\boldsymbol{Z}}_{j}\right)=\hat{\lambda}_{j}^{*},\quad j=1,\cdots,p,\\&\operatorname{cov}\left(\tilde{\boldsymbol{Z}}_{i},\tilde{\boldsymbol{Z}}_{j}\right)=0,\quad i\neq j,\quad i,j=1,\cdots,p.\end{aligned}$$

(R3).

$$\operatorname{cov}\left(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Z}}\right)=\left(\operatorname{cov}\left(\tilde{\boldsymbol{X}}_{i},\tilde{\boldsymbol{Z}}_{j}\right)\right)_{p\times p}=\left(\hat{q}_{ij}^{*}\hat{\lambda}_{j}^{*}\right)_{p\times p}=\hat{\boldsymbol{Q}}^{*}\hat{\Lambda}^{*},$$

特别地，

$$\operatorname{cov}\left(\tilde{\boldsymbol{X}}_{i},\tilde{\boldsymbol{Z}}_{j}\right)=\hat{q}_{ij}^{*}\hat{\lambda}_{j}^{*},\quad i,j=1,\cdots,p.$$

(R4).

$$\begin{aligned}\mathrm{cor}\left(\tilde{\boldsymbol{X}}\right)&=\mathrm{cor}\left(\boldsymbol{X}^*\right)=\mathrm{cov}\left(\boldsymbol{X}^{**}\right)=\mathrm{cov}\left(\boldsymbol{X}^*\right)=\mathrm{cor}\left(\boldsymbol{X}\right)=\boldsymbol{R}=\left(r_{ij}\right)_{p\times p}\\&=\mathrm{cor}\left(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{X}}\right)=\left(\mathrm{cor}\left(\tilde{\boldsymbol{X}}_i,\tilde{\boldsymbol{X}}_j\right)\right)_{p\times p},\end{aligned}$$

特别地，对 $i,j=1,\cdots,p,$

$$\begin{aligned}\mathrm{cor}\left(\tilde{\boldsymbol{X}}_i,\tilde{\boldsymbol{X}}_j\right)&=\mathrm{cor}\left(\boldsymbol{X}_i^*,\boldsymbol{X}_j^*\right)=\mathrm{cov}\left(\boldsymbol{X}_i^{**},\boldsymbol{X}_j^{**}\right)=\mathrm{cov}\left(\boldsymbol{X}_i^*,\boldsymbol{X}_j^*\right)\\&=\mathrm{cov}\left(\tilde{\boldsymbol{X}}_i,\tilde{\boldsymbol{X}}_j\right)=\mathrm{cor}\left(\boldsymbol{X}_i,\boldsymbol{X}_j\right)=r_{ij}.\end{aligned}$$

(R5). 若 $\boldsymbol{R}>\mathbf{0}$，则

$$\begin{aligned}\mathrm{cor}\left(\tilde{\boldsymbol{Z}}\right)&=\mathrm{cor}\left(\tilde{\boldsymbol{Z}},\tilde{\boldsymbol{Z}}\right)=\left(\mathrm{cor}\left(\tilde{\boldsymbol{Z}}_i,\tilde{\boldsymbol{Z}}_j\right)\right)_{p\times p}=\mathrm{cov}\left(\tilde{\boldsymbol{Z}}^*\right)\\&=\boldsymbol{E}_{p\times p}=\mathrm{diag}\left(1,1,\cdots,1\right)_{p\times p},\end{aligned}$$

特别地，

$$\begin{aligned}&\mathrm{cor}\left(\tilde{\boldsymbol{Z}}_j\right)=\mathrm{cor}\left(\tilde{\boldsymbol{Z}}_j,\tilde{\boldsymbol{Z}}_j\right)=1,\quad j=1,\cdots,p,\\&\mathrm{cor}\left(\tilde{\boldsymbol{Z}}_i,\tilde{\boldsymbol{Z}}_j\right)=0,\quad i\neq j,\quad i,j=1,\cdots,p.\end{aligned}$$

(R6). 若 $\boldsymbol{R}>\mathbf{0}$，则

$$\begin{aligned}\mathrm{cor}\left(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Z}}\right)&=\left(\mathrm{cor}\left(\tilde{\boldsymbol{X}}_i,\tilde{\boldsymbol{Z}}_j\right)\right)_{p\times p}=\mathrm{cov}\left(\tilde{\boldsymbol{X}}^*,\tilde{\boldsymbol{Z}}^*\right)=\left(\mathrm{cov}\left(\boldsymbol{X}_i^*,\tilde{\boldsymbol{Z}}_j^*\right)\right)_{p\times p}\\&=\hat{\boldsymbol{Q}}^*\left(\hat{\Lambda}^*\right)^{\frac{1}{2}}=\left(\hat{q}_{ij}^*\sqrt{\hat{\lambda}_j^*}\right)_{p\times p},\end{aligned}$$

特别地，

$$\mathrm{cor}\left(\tilde{\boldsymbol{X}}_i,\tilde{\boldsymbol{Z}}_j\right)=\mathrm{cov}\left(\boldsymbol{X}_i^*,\tilde{\boldsymbol{Z}}_j^*\right)=\hat{q}_{ij}^*\sqrt{\hat{\lambda}_j^*},\quad i,j=1,\cdots,p.$$

(R7). 样本总方差

$$\sum_{j=1}^{p}\mathrm{cov}\left(\tilde{\boldsymbol{X}}_j\right)=\sum_{j=1}^{p}1=p=\sum_{j=1}^{p}\hat{\lambda}_j^*=\sum_{j=1}^{p}\mathrm{cov}\left(\tilde{\boldsymbol{Z}}_j\right).$$

此式表明了样本 $\tilde{\boldsymbol{X}}_1,\cdots,\tilde{\boldsymbol{X}}_p$ 的样本(协)方差之和等于样本主成分 $\tilde{\boldsymbol{Z}}_1,\cdots,\tilde{\boldsymbol{Z}}_p$ 的样本(协)方差之和。

在以上 7 个性质中，(R1)显然，(R2)和(R6)的分量形式的结果和(R7)可以在[7]中找到。

注意，在(R5)和(R6)中我们要求 $\boldsymbol{R}>\mathbf{0}$，而其余性质只要求 $\boldsymbol{R}\geq\mathbf{0}$。因为在(R5)和(R6)的证明中会涉及到

$$\left(\hat{\Lambda}^*\right)^{-\frac{1}{2}}=\mathrm{diag}\left(\frac{1}{\sqrt{\hat{\lambda}_1^*}},\frac{1}{\sqrt{\hat{\lambda}_2^*}},\cdots,\frac{1}{\sqrt{\hat{\lambda}_p^*}}\right),$$

若 $\boldsymbol{R}$ 有一个 0 特征值，即 $\hat{\lambda}_p^*=0$，则 $1\big/\sqrt{\hat{\lambda}_p^*}$ 是没有意义的，从而结论不成立，数值结果也证实了这一点。

由性质(R1)-(R6)归纳得到样本 $\tilde{\boldsymbol{X}}_i,\tilde{\boldsymbol{X}}_j,\tilde{\boldsymbol{Z}}_i,\tilde{\boldsymbol{Z}}_j$ 的样本协方差关系图((R1)-(R3))和样本相关系数关系图((R4)-(R6))，如图 2。

性质(R1)-(R7)的证明有些是初等的，有些需要一定的技巧，由于这些证明较占篇幅，所以把它们放到了本文的附录中。

3. 数值模拟

此节我们在 R 软件[11]中编写程序，通过数值模拟来验证从样本协方差矩阵 $\boldsymbol{S}$ 出发计算的样本主成分的性质(S1)-(S7)及从样本相关矩阵 $\boldsymbol{R}$ 出发计算的样本主成分的性质(R1)-(R7)。

在 R 软件中，求矩阵的样本协方差矩阵的函数是 cov()，求矩阵的样本相关矩阵的函数是 cor() [12]。性质(S1)-(S7)即是要验证

$$\text{(S1).}\ \ \operatorname{cov}(\boldsymbol{X})=\boldsymbol{S},\ \ \text{(S2).}\ \ \operatorname{cov}(\boldsymbol{Z})=\hat{\Lambda},\ \ \text{(S3).}\ \ \operatorname{cov}(\boldsymbol{X},\boldsymbol{Z})=\hat{\boldsymbol{Q}}\hat{\Lambda},$$

$$\text{(S4).}\ \ \operatorname{cor}(\boldsymbol{X})=\boldsymbol{R},\ \ \text{(S5).}\ \ \operatorname{cor}(\boldsymbol{Z})=\boldsymbol{E}_{p\times p},$$

$$\text{(S6).}\ \ \operatorname{cor}(\boldsymbol{X},\boldsymbol{Z})=\hat{\boldsymbol{D}}^{-\frac{1}{2}}\hat{\boldsymbol{Q}}\hat{\Lambda}^{\frac{1}{2}}=\left(\operatorname{Diag}(\boldsymbol{S})\right)^{-\frac{1}{2}}\hat{\boldsymbol{Q}}\hat{\Lambda}^{\frac{1}{2}},$$

$$\text{(S7).}\ \ \operatorname{tr}\left(\operatorname{cov}(\boldsymbol{X})\right)=\operatorname{tr}\left(\operatorname{cov}(\boldsymbol{Z})\right).$$

注意在(S1)中的 $\boldsymbol{S}$ 和(S4)中的 $\boldsymbol{R}$ 分别为

$$\boldsymbol{S}=\frac{1}{n-1}\sum_{k=1}^{n}\left(\boldsymbol{X}_{(k)}-\bar{\boldsymbol{X}}\right)\left(\boldsymbol{X}_{(k)}-\bar{\boldsymbol{X}}\right)^{\mathrm{T}},$$

$$\boldsymbol{R}=\frac{1}{n-1}\sum_{k=1}^{n}\left(\boldsymbol{X}_{(k)}^{*}-\overline{\boldsymbol{X}^{*}}\right)\left(\boldsymbol{X}_{(k)}^{*}-\overline{\boldsymbol{X}^{*}}\right)^{\mathrm{T}}.$$

性质(R1)-(R7)即是要验证

$$\text{(R1).}\ \ \operatorname{cov}(\tilde{\boldsymbol{X}})=\boldsymbol{R},\ \ \text{(R2).}\ \ \operatorname{cov}(\tilde{\boldsymbol{Z}})=\hat{\Lambda}^{*},\ \ \text{(R3).}\ \ \operatorname{cov}(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Z}})=\hat{\boldsymbol{Q}}^{*}\hat{\Lambda}^{*},$$

$$\text{(R4).}\ \ \operatorname{cor}(\tilde{\boldsymbol{X}})=\boldsymbol{R},\ \ \text{(R5).}\ \ \operatorname{cor}(\tilde{\boldsymbol{Z}})=\boldsymbol{E}_{p\times p},$$

$$\text{(R6).}\ \ \operatorname{cor}(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Z}})=\hat{\boldsymbol{Q}}^{*}\left(\hat{\Lambda}^{*}\right)^{\frac{1}{2}},\ \ \text{(R7).}\ \ \operatorname{tr}\left(\operatorname{cov}(\tilde{\boldsymbol{X}})\right)=p=\operatorname{tr}\left(\operatorname{cov}(\tilde{\boldsymbol{Z}})\right).$$

注意在(S5)和(S6)中要求 $\boldsymbol{S}>\boldsymbol{0}$，在(R5)和(R6)中要求 $\boldsymbol{R}>\boldsymbol{0}$。

下面我们举两个例子来验证(S1)-(S7)和(R1)-(R7)的正确性。

例 1.

$$\boldsymbol{X}=\left(x_{ij}\right)_{n\times p},\ n=4,\ p=3,\ x_{ij}\sim N(0,1).$$

为重复本文的结果，我们在 R 软件中使用 set.seed(1)，此时

$$\boldsymbol{X}=\begin{pmatrix}-0.6264538 & 0.3295078 & 0.5757814\\ 0.1836433 & -0.8204684 & -0.3053884\\ -0.8356286 & 0.4874291? & 1.5117812\\ 1.5952808 & 0.7383247 & 0.3898432\end{pmatrix}.$$

容易验证 $\boldsymbol{S}>\boldsymbol{0}$ 且 $\boldsymbol{R}>\boldsymbol{0}$，(S1)-(S7)和(R1)-(R7)均是正确的。详细的数值模拟结果请见本文的附录。

下面给一个 $\boldsymbol{S}\geq\boldsymbol{0}$ 且 $\boldsymbol{R}\geq\boldsymbol{0}$ 的例子，即 $\boldsymbol{S}$ 和 $\boldsymbol{R}$ 都有 1 个 0 特征值，此时(S5)，(S6)，(R5)和(R6)是不成立的，但其余性质是成立的。

例 2.

$$\boldsymbol{X}=\left(x_{ij}\right)_{n\times p},\ n=p=3,\ x_{ij}\sim N(0,1).$$

为重复本文的结果，我们在 R 软件中使用 set.seed(1)，此时

$$\boldsymbol{X}=\begin{pmatrix}-0.6264538 & 1.5952808 & 0.4874291\\ 0.1836433 & 0.3295078 & 0.7383247\\ -0.8356286 & -0.8204684 & 0.5757814\end{pmatrix}.$$

详细的数值模拟结果请见本文的附录。

4. 总结

我们给出了样本主成分的性质及证明，分两种情况讨论：从 $\boldsymbol{S}$ 出发求主成分和从 $\boldsymbol{R}$ 出发求主成分。在从 $\boldsymbol{S}$ 出发求主成分中，给出了 7 个性质(S1)-(S7)及它们的证明，这些性质说明的关系在图 1 中得到了充分的展现。同样，在从 $\boldsymbol{R}$ 出发求主成分中，给出了 7 个性质(R1)-(R7)及它们的证明，这些性质说明的关系在图 2 中得到了充分的展现。最后我们给出了两个数值模拟的例子来验证性质(S1)-(S7)和(R1)-(R7)的正确性。例 1 中的 $\boldsymbol{S}>\boldsymbol{0}$ 且 $\boldsymbol{R}>\boldsymbol{0}$，从而性质(S1)-(S7)和(R1)-(R7)均是正确的。例 2 中的 $\boldsymbol{S}\geq\boldsymbol{0}$ 且 $\boldsymbol{R}\geq\boldsymbol{0}$，即 $\boldsymbol{S}$ 和 $\boldsymbol{R}$ 都有 1 个 0 特征值，数值模拟显示(S5)，(S6)，(R5)和(R6)是不成立的，但其余性质是成立的，与理论结果相一致。

基金项目

本研究受中央高校基本科研业务费项目(CQDXWL-2012-004; CDJRC10100010)，中国国家留学基金(201606055028)和教育部人文社会科学研究西部和边疆地区项目(14XJC910001)支持。

参考文献 (References)

[1] 张应应. 总体或样本的协方差(矩阵)和相关系数(矩阵)的系统定义[J]. 统计与决策, 2016(8): 20-24.

[2] 杨虎, 刘琼荪, 钟波. 数理统计[M]. 北京: 高等教育出版社, 2004.

[3] 高惠璇. 应用多元统计分析[M]. 北京: 北京大学出版社, 2005.

[4] Johnson, R.A. and Wichern, D.W. (2007) Applied Multivariate Statistical Analysis. 6th Edition, Pearson, Hoboken.

[5] 汤银才. R 语言与统计分析[M]. 北京: 高等教育出版社, 2008.

[6] 王学民. 应用多元分析[M]. 第 3 版. 上海: 上海财经大学出版社, 2009.

[7] 薛毅, 陈立萍. 统计建模与 R 软件[M]. 北京: 清华大学出版社, 2009.

[8] 任雪松, 于秀林. 多元统计分析[M]. 第 2 版. 北京: 中国统计出版社, 2011.

[9] 何晓群. 多元统计分析[M]. 第 3 版. 北京: 中国人民大学出版社, 2012.

[10] 李素兰. 数据分析与 R 软件[M]. 北京: 科学出版社, 2013.

[11] R Core Team (2017) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/

[12] Becker, R.A., Chambers, J.M. and Wilks, A.R. (1988) The New S Language. Wadsworth & Brooks/Cole.

附　录

附录中包含了性质(S1)-(S7)和(R1)-(R7)的证明，还有两个例子的详细的数值模拟结果。

A.1. 准备知识及杂项证明

准备知识

样本的协方差(矩阵)和相关系数(矩阵)的定义及性质请参见[1]。[1]中有三个公式在本文中经常用到：

$$\mathrm{cov}(\boldsymbol{X},\boldsymbol{Y})=\left(\mathrm{cov}(\boldsymbol{X}_i,\boldsymbol{Y}_j)\right)_{p\times q}, \tag{A.1}$$

$$\mathrm{cor}(\boldsymbol{X},\boldsymbol{Y})=\left[\mathrm{Diag}\left(\mathrm{cov}(\boldsymbol{X})\right)\right]^{-\frac{1}{2}}\mathrm{cov}(\boldsymbol{X},\boldsymbol{Y})\left[\mathrm{Diag}\left(\mathrm{cov}(\boldsymbol{Y})\right)\right]^{-\frac{1}{2}}, \tag{A.2}$$

其中 $\mathrm{Diag}\left(\mathrm{cov}(\boldsymbol{X})\right)$ 表示取矩阵 $\mathrm{cov}(\boldsymbol{X})$ 的对角线元素构成的对角阵，

$$\mathrm{cor}(\boldsymbol{X},\boldsymbol{Y})=\left(\mathrm{cor}(\boldsymbol{X}_i,\boldsymbol{Y}_j)\right)_{p\times q}. \tag{A.3}$$

另外，还有

$$\boldsymbol{X}^{**}=\boldsymbol{X}^{*}=\left(x_{ij}^{*}\right)_{n\times p}. \tag{A.4}$$

杂项证明

因为 $\boldsymbol{Z}_{(k)}=\hat{\boldsymbol{Q}}^{\mathrm{T}}\boldsymbol{X}_{(k)}$，人们自然会想到 $\boldsymbol{Z}_{(k)}^{*}=\hat{\boldsymbol{Q}}^{\mathrm{T}}\boldsymbol{X}_{(k)}^{*}$，但是一般来说这是错误的。从[1]中，我们知道

$$\boldsymbol{X}_{(k)}^{*}=\hat{\boldsymbol{D}}^{-\frac{1}{2}}\left(\boldsymbol{X}_{(k)}-\bar{\boldsymbol{X}}\right),$$

其中，

$$\hat{\boldsymbol{D}}=\hat{\boldsymbol{D}}^{X}=\mathrm{diag}\left(s_{11},s_{22},\cdots,s_{pp}\right)=\mathrm{Diag}(\boldsymbol{S})=\mathrm{Diag}\left(\mathrm{cov}(\boldsymbol{X})\right).$$

易知

$$\hat{\boldsymbol{Q}}^{\mathrm{T}}\boldsymbol{X}_{(k)}^{*}=\hat{\boldsymbol{Q}}^{\mathrm{T}}\hat{\boldsymbol{D}}^{-\frac{1}{2}}\left(\boldsymbol{X}_{(k)}-\bar{\boldsymbol{X}}\right),$$

$$\boldsymbol{Z}_{(k)}^{*}=\hat{\boldsymbol{\Lambda}}^{-\frac{1}{2}}\left(\boldsymbol{Z}_{(k)}-\bar{\boldsymbol{Z}}\right)=\hat{\boldsymbol{\Lambda}}^{-\frac{1}{2}}\left(\hat{\boldsymbol{Q}}^{\mathrm{T}}\boldsymbol{X}_{(k)}-\hat{\boldsymbol{Q}}^{\mathrm{T}}\bar{\boldsymbol{X}}\right)=\hat{\boldsymbol{\Lambda}}^{-\frac{1}{2}}\hat{\boldsymbol{Q}}^{\mathrm{T}}\left(\boldsymbol{X}_{(k)}-\bar{\boldsymbol{X}}\right).$$

则

$$\hat{\boldsymbol{Q}}^{\mathrm{T}}\boldsymbol{X}_{(k)}^{*}=\boldsymbol{Z}_{(k)}^{*}\Leftrightarrow\hat{\boldsymbol{Q}}^{\mathrm{T}}\hat{\boldsymbol{D}}^{-\frac{1}{2}}=\hat{\Lambda}^{-\frac{1}{2}}\hat{\boldsymbol{Q}}^{\mathrm{T}}$$

$$\Leftrightarrow\begin{pmatrix}\hat{q}_{11}&\hat{q}_{21}&\cdots&\hat{q}_{p1}\\\hat{q}_{12}&\hat{q}_{22}&\cdots&\hat{q}_{p2}\\\vdots&\vdots&&\vdots\\\hat{q}_{1p}&\hat{q}_{2p}&\cdots&\hat{q}_{pp}\end{pmatrix}\begin{pmatrix}\frac{1}{\sqrt{s_{11}}}&&&\\&\frac{1}{\sqrt{s_{22}}}&&\\&&\ddots&\\&&&\frac{1}{\sqrt{s_{pp}}}\end{pmatrix}$$

$$=\begin{pmatrix}\frac{1}{\sqrt{\hat{\lambda}_1}}&&&\\&\frac{1}{\sqrt{\hat{\lambda}_2}}&&\\&&\ddots&\\&&&\frac{1}{\sqrt{\hat{\lambda}_p}}\end{pmatrix}\begin{pmatrix}\hat{q}_{11}&\hat{q}_{21}&\cdots&\hat{q}_{p1}\\\hat{q}_{12}&\hat{q}_{22}&\cdots&\hat{q}_{p2}\\\vdots&\vdots&&\vdots\\\hat{q}_{1p}&\hat{q}_{2p}&\cdots&\hat{q}_{pp}\end{pmatrix}$$

$$\Leftrightarrow\begin{pmatrix}\frac{\hat{q}_{11}}{\sqrt{s_{11}}}&\frac{\hat{q}_{21}}{\sqrt{s_{22}}}&\cdots&\frac{\hat{q}_{p1}}{\sqrt{s_{pp}}}\\\frac{\hat{q}_{12}}{\sqrt{s_{11}}}&\frac{\hat{q}_{22}}{\sqrt{s_{22}}}&\cdots&\frac{\hat{q}_{p2}}{\sqrt{s_{pp}}}\\\vdots&\vdots&&\vdots\\\frac{\hat{q}_{1p}}{\sqrt{s_{11}}}&\frac{\hat{q}_{2p}}{\sqrt{s_{22}}}&\cdots&\frac{\hat{q}_{pp}}{\sqrt{s_{pp}}}\end{pmatrix}=\begin{pmatrix}\frac{\hat{q}_{11}}{\sqrt{\hat{\lambda}_1}}&\frac{\hat{q}_{21}}{\sqrt{\hat{\lambda}_1}}&\cdots&\frac{\hat{q}_{p1}}{\sqrt{\hat{\lambda}_1}}\\\frac{\hat{q}_{12}}{\sqrt{\hat{\lambda}_2}}&\frac{\hat{q}_{22}}{\sqrt{\hat{\lambda}_2}}&\cdots&\frac{\hat{q}_{p2}}{\sqrt{\hat{\lambda}_2}}\\\vdots&\vdots&&\vdots\\\frac{\hat{q}_{1p}}{\sqrt{\hat{\lambda}_p}}&\frac{\hat{q}_{2p}}{\sqrt{\hat{\lambda}_p}}&\cdots&\frac{\hat{q}_{pp}}{\sqrt{\hat{\lambda}_p}}\end{pmatrix}$$

$$\Leftrightarrow s_{11}=s_{22}=\cdots=s_{pp}=\hat{\lambda}_1=\hat{\lambda}_2=\cdots=\hat{\lambda}_p.$$

上式一般来说是不会一定成立的，所以一般来说 $\boldsymbol{Z}_{(k)}^{*}\neq\hat{\boldsymbol{Q}}^{\mathrm{T}}\boldsymbol{X}_{(k)}^{*}$。

因为 $\tilde{\boldsymbol{Z}}_{(k)}=\hat{\boldsymbol{Q}}^{*\mathrm{T}}\tilde{\boldsymbol{X}}_{(k)}$，人们自然会想到 $\tilde{\boldsymbol{Z}}_{(k)}^{*}=\hat{\boldsymbol{Q}}^{*\mathrm{T}}\tilde{\boldsymbol{X}}_{(k)}^{*}$，同样，一般来说这是错误的。我们有

$$\hat{\boldsymbol{Q}}^{*\mathrm{T}}\tilde{\boldsymbol{X}}_{(k)}^{*}=\hat{\boldsymbol{Q}}^{*\mathrm{T}}\boldsymbol{X}_{(k)}^{**}=\hat{\boldsymbol{Q}}^{*\mathrm{T}}\boldsymbol{X}_{(k)}^{*}=\hat{\boldsymbol{Q}}^{*\mathrm{T}}\tilde{\boldsymbol{X}}_{(k)},$$

$$\tilde{\boldsymbol{Z}}_{(k)}^{*}=\left(\hat{\Lambda}^{*}\right)^{-\frac{1}{2}}\tilde{\boldsymbol{Z}}_{(k)}=\left(\hat{\Lambda}^{*}\right)^{-\frac{1}{2}}\hat{\boldsymbol{Q}}^{*\mathrm{T}}\tilde{\boldsymbol{X}}_{(k)}.$$

则

$$\hat{\boldsymbol{Q}}^{*\mathrm{T}}\tilde{\boldsymbol{X}}_{(k)}^{*}=\tilde{\boldsymbol{Z}}_{(k)}^{*}\Leftrightarrow\left(\hat{\Lambda}^{*}\right)^{-\frac{1}{2}}=\boldsymbol{E}_{p\times p}\Leftrightarrow\hat{\Lambda}^{*}=\boldsymbol{E}_{p\times p}.$$

同样，上式一般来说是不会一定成立的，所以一般来说 $\tilde{\boldsymbol{Z}}_{(k)}^{*}\neq\hat{\boldsymbol{Q}}^{*\mathrm{T}}\tilde{\boldsymbol{X}}_{(k)}^{*}$。

A.2. 从 S 出发求主成分

证明：

(S1). 由数据资料阵 $\boldsymbol{X}$ 的样本协方差矩阵的定义有

$$\mathrm{cov}(\boldsymbol{X})=\mathrm{cov}(\boldsymbol{X},\boldsymbol{X})=\boldsymbol{S}^{X}=\boldsymbol{S}=(s_{ij})_{p\times p}.$$

由(A.1)得，

$$\mathrm{cov}(\boldsymbol{X},\boldsymbol{X})=\left(\mathrm{cov}(\boldsymbol{X}_i,\boldsymbol{X}_j)\right)_{p\times p}.$$

从而，

$$\mathrm{cov}(\boldsymbol{X}_i,\boldsymbol{X}_j)=s_{ij}=\frac{1}{n-1}\sum_{k=1}^{n}(x_{ki}-\bar{x}_i)(x_{kj}-\bar{x}_j),\ \ i,j=1,\cdots,p.$$

(S2). 易知

$$\boldsymbol{Z}_{(k)}-\bar{\boldsymbol{Z}}=\hat{\boldsymbol{Q}}^{\mathrm{T}}\boldsymbol{X}_{(k)}-\hat{\boldsymbol{Q}}^{\mathrm{T}}\bar{\boldsymbol{X}}=\hat{\boldsymbol{Q}}^{\mathrm{T}}\left(\boldsymbol{X}_{(k)}-\bar{\boldsymbol{X}}\right),$$

从而由矩阵 $\boldsymbol{Z}$ 的样本协方差矩阵的定义有

$$\begin{aligned}\mathrm{cov}(\boldsymbol{Z})=\mathrm{cov}(\boldsymbol{Z},\boldsymbol{Z})&=\frac{1}{n-1}\sum_{k=1}^{n}\left(\boldsymbol{Z}_{(k)}-\bar{\boldsymbol{Z}}\right)\left(\boldsymbol{Z}_{(k)}-\bar{\boldsymbol{Z}}\right)^{\mathrm{T}}\\&=\frac{1}{n-1}\sum_{k=1}^{n}\hat{\boldsymbol{Q}}^{\mathrm{T}}\left(\boldsymbol{X}_{(k)}-\bar{\boldsymbol{X}}\right)\left(\boldsymbol{X}_{(k)}-\bar{\boldsymbol{X}}\right)^{\mathrm{T}}\hat{\boldsymbol{Q}}\\&=\hat{\boldsymbol{Q}}^{\mathrm{T}}\left[\frac{1}{n-1}\sum_{k=1}^{n}\left(\boldsymbol{X}_{(k)}-\bar{\boldsymbol{X}}\right)\left(\boldsymbol{X}_{(k)}-\bar{\boldsymbol{X}}\right)^{\mathrm{T}}\right]\hat{\boldsymbol{Q}}\\&=\hat{\boldsymbol{Q}}^{\mathrm{T}}\boldsymbol{S}\hat{\boldsymbol{Q}}=\hat{\Lambda}.\end{aligned}$$

由(A.1)得

$$\mathrm{cov}(\boldsymbol{Z},\boldsymbol{Z})=\left(\mathrm{cov}(\boldsymbol{Z}_i,\boldsymbol{Z}_j)\right)_{p\times p}=\hat{\Lambda}=\begin{pmatrix}\hat{\lambda}_1&&&\\&\hat{\lambda}_2&&\\&&\ddots&\\&&&\hat{\lambda}_p\end{pmatrix}.$$

从而

$$\mathrm{cov}(\boldsymbol{Z}_j)=\mathrm{cov}(\boldsymbol{Z}_j,\boldsymbol{Z}_j)=\hat{\lambda}_j,\quad j=1,\cdots,p,$$
$$\mathrm{cov}(\boldsymbol{Z}_i,\boldsymbol{Z}_j)=0,\quad i\neq j,\quad i,j=1,\cdots,p.$$

在证明 $\mathrm{cov}(\boldsymbol{Z}_j)=\hat{\lambda}_j$ 时，我们也可以利用样本协方差的定义直接来求。

$$\boldsymbol{Z}_j=\begin{pmatrix}z_{1j}\\z_{2j}\\\vdots\\z_{nj}\end{pmatrix}=\mathrm{X}\hat{\boldsymbol{a}}_j=\begin{pmatrix}\boldsymbol{X}_{(1)}^{\mathrm{T}}\\\boldsymbol{X}_{(2)}^{\mathrm{T}}\\\vdots\\\boldsymbol{X}_{(n)}^{\mathrm{T}}\end{pmatrix}\hat{\boldsymbol{a}}_j=\begin{pmatrix}\boldsymbol{X}_{(1)}^{\mathrm{T}}\hat{\boldsymbol{a}}_j\\\boldsymbol{X}_{(2)}^{\mathrm{T}}\hat{\boldsymbol{a}}_j\\\vdots\\\boldsymbol{X}_{(n)}^{\mathrm{T}}\hat{\boldsymbol{a}}_j\end{pmatrix}=\begin{pmatrix}\hat{\boldsymbol{a}}_j^{\mathrm{T}}\boldsymbol{X}_{(1)}\\\hat{\boldsymbol{a}}_j^{\mathrm{T}}\boldsymbol{X}_{(2)}\\\vdots\\\hat{\boldsymbol{a}}_j^{\mathrm{T}}\boldsymbol{X}_{(n)}\end{pmatrix},$$

从而

$$z_{kj}=\hat{\boldsymbol{a}}_j^{T}\boldsymbol{X}_{(k)},\quad k=1,\cdots,n,$$
$$\bar{z}_j=\frac{1}{n}\sum_{k=1}^{n}z_{kj}=\frac{1}{n}\sum_{k=1}^{n}\hat{\boldsymbol{a}}_j^{\mathrm{T}}\boldsymbol{X}_{(k)}=\hat{\boldsymbol{a}}_j^{\mathrm{T}}\left(\frac{1}{n}\sum_{k=1}^{n}\boldsymbol{X}_{(k)}\right)=\hat{\boldsymbol{a}}_j^{\mathrm{T}}\bar{\boldsymbol{X}},$$
$$z_{kj}-\bar{z}_j=\hat{\boldsymbol{a}}_j^{\mathrm{T}}\boldsymbol{X}_{(k)}-\hat{\boldsymbol{a}}_j^{\mathrm{T}}\bar{\boldsymbol{X}}=\hat{\boldsymbol{a}}_j^{\mathrm{T}}\left(\boldsymbol{X}_{(k)}-\bar{\boldsymbol{X}}\right)=\left(\boldsymbol{X}_{(k)}-\bar{\boldsymbol{X}}\right)^{\mathrm{T}}\hat{\boldsymbol{a}}_j.$$

因此

$$\begin{aligned}\operatorname{cov}\left(\boldsymbol{Z}_j\right)&=\operatorname{cov}\left(\boldsymbol{Z}_j,\boldsymbol{Z}_j\right)\\&=\frac{1}{n-1}\sum_{k=1}^{n}\left(z_{kj}-\overline{z}_j\right)^2\\&=\frac{1}{n-1}\sum_{k=1}^{n}\hat{\boldsymbol{a}}_j^{\mathrm{T}}\left(\boldsymbol{X}_{(k)}-\overline{\boldsymbol{X}}\right)\left(\boldsymbol{X}_{(k)}-\overline{\boldsymbol{X}}\right)^{\mathrm{T}}\hat{\boldsymbol{a}}_j\\&=\hat{\boldsymbol{a}}_j^{T}\left[\frac{1}{n-1}\sum_{k=1}^{n}\left(\boldsymbol{X}_{(k)}-\overline{\boldsymbol{X}}\right)\left(\boldsymbol{X}_{(k)}-\overline{\boldsymbol{X}}\right)^{\mathrm{T}}\right]\hat{\boldsymbol{a}}_j\\&=\hat{\boldsymbol{a}}_j^{\mathrm{T}}\boldsymbol{S}\hat{\boldsymbol{a}}_j=\hat{\boldsymbol{a}}_j^{\mathrm{T}}\hat{\lambda}_j\hat{\boldsymbol{a}}_j=\hat{\lambda}_j.\end{aligned}$$

类似地，在证明 $\operatorname{cov}\left(\boldsymbol{Z}_i,\boldsymbol{Z}_j\right)=0, i\neq j, i, j=1,\cdots,p$ 时，我们也可以利用样本协方差的定义直接来求。

$$\boldsymbol{Z}_i=\begin{pmatrix}z_{1i}\\z_{2i}\\\vdots\\z_{ni}\end{pmatrix}=\boldsymbol{X}\hat{\boldsymbol{a}}_i=\begin{pmatrix}\boldsymbol{X}_{(1)}^{\mathrm{T}}\\\boldsymbol{X}_{(2)}^{\mathrm{T}}\\\vdots\\\boldsymbol{X}_{(n)}^{\mathrm{T}}\end{pmatrix}\hat{\boldsymbol{a}}_i=\begin{pmatrix}\boldsymbol{X}_{(1)}^{\mathrm{T}}\hat{\boldsymbol{a}}_i\\\boldsymbol{X}_{(2)}^{\mathrm{T}}\hat{\boldsymbol{a}}_i\\\vdots\\\boldsymbol{X}_{(n)}^{\mathrm{T}}\hat{\boldsymbol{a}}_i\end{pmatrix}=\begin{pmatrix}\hat{\boldsymbol{a}}_i^{\mathrm{T}}\boldsymbol{X}_{(1)}\\\hat{\boldsymbol{a}}_i^{\mathrm{T}}\boldsymbol{X}_{(2)}\\\vdots\\\hat{\boldsymbol{a}}_i^{\mathrm{T}}\boldsymbol{X}_{(n)}\end{pmatrix},$$

$$z_{ki}-\overline{z}_i=\hat{\boldsymbol{a}}_i^{\mathrm{T}}\boldsymbol{X}_{(k)}-\hat{\boldsymbol{a}}_i^{\mathrm{T}}\overline{\boldsymbol{X}}=\hat{\boldsymbol{a}}_i^{\mathrm{T}}\left(\boldsymbol{X}_{(k)}-\overline{\boldsymbol{X}}\right)=\left(\boldsymbol{X}_{(k)}-\overline{\boldsymbol{X}}\right)^{\mathrm{T}}\hat{\boldsymbol{a}}_i.$$

因此对 $i\neq j, i, j=1,\cdots,p$,

$$\begin{aligned}\operatorname{cov}\left(\boldsymbol{Z}_i,\boldsymbol{Z}_j\right)&=\frac{1}{n-1}\sum_{k=1}^{n}\left(z_{ki}-\overline{z}_i\right)\left(z_{kj}-\overline{z}_j\right)\\&=\frac{1}{n-1}\sum_{k=1}^{n}\hat{\boldsymbol{a}}_i^{\mathrm{T}}\left(\boldsymbol{X}_{(k)}-\overline{\boldsymbol{X}}\right)\left(\boldsymbol{X}_{(k)}-\overline{\boldsymbol{X}}\right)^{\mathrm{T}}\hat{\boldsymbol{a}}_j\\&=\hat{\boldsymbol{a}}_i^{\mathrm{T}}\left[\frac{1}{n-1}\sum_{k=1}^{n}\left(\boldsymbol{X}_{(k)}-\overline{\boldsymbol{X}}\right)\left(\boldsymbol{X}_{(k)}-\overline{\boldsymbol{X}}\right)^{\mathrm{T}}\right]\hat{\boldsymbol{a}}_j\\&=\hat{\boldsymbol{a}}_i^{\mathrm{T}}\boldsymbol{S}\hat{\boldsymbol{a}}_j=\hat{\boldsymbol{a}}_i^{\mathrm{T}}\hat{\lambda}_j\hat{\boldsymbol{a}}_j=\hat{\lambda}_j\hat{\boldsymbol{a}}_i^{\mathrm{T}}\hat{\boldsymbol{a}}_j=0.\end{aligned}$$

由此我们发现用矩阵形式来证明比用分量形式来证明要简洁得多。

(S3). 由(A.1)，得

$$\operatorname{cov}\left(\boldsymbol{X},\boldsymbol{Z}\right)=\left(\operatorname{cov}\left(\boldsymbol{X}_i,\boldsymbol{Z}_j\right)\right)_{p\times p}.$$

由 $\boldsymbol{Z}_{(k)}=\hat{\boldsymbol{Q}}^{\mathrm{T}}\boldsymbol{X}_{(k)}$，得

$$\boldsymbol{X}_{(k)}=\hat{\boldsymbol{Q}}\boldsymbol{Z}_{(k)},\tag{A.5}$$

即

$$\begin{pmatrix}x_{k1}\\x_{k2}\\\vdots\\x_{kp}\end{pmatrix}=\left(\hat{q}_{ij}\right)_{p\times p}\begin{pmatrix}z_{k1}\\z_{k2}\\\vdots\\z_{kp}\end{pmatrix},$$

则

$$x_{ki}=\left(\hat{q}_{i1},\hat{q}_{i2},\cdots,\hat{q}_{ip}\right)\begin{pmatrix}z_{k1}\\z_{k2}\\\vdots\\z_{kp}\end{pmatrix}=\sum_{r=1}^{p}\hat{q}_{ir}z_{kr}=\sum_{r=1}^{p}z_{kr}\hat{q}_{ir},$$

$$\begin{aligned}\bar{x}_i&=\frac{1}{n}\sum_{k=1}^{n}x_{ki}=\frac{1}{n}\sum_{k=1}^{n}\sum_{r=1}^{p}z_{kr}\hat{q}_{ir}=\frac{1}{n}\sum_{r=1}^{p}\sum_{k=1}^{n}z_{kr}\hat{q}_{ir}\\&=\sum_{r=1}^{p}\left(\frac{1}{n}\sum_{k=1}^{n}z_{kr}\right)\hat{q}_{ir}=\sum_{r=1}^{p}\bar{z}_r\hat{q}_{ir},\end{aligned}$$

$$x_{ki}-\bar{x}_i=\sum_{r=1}^{p}z_{kr}\hat{q}_{ir}-\sum_{r=1}^{p}\bar{z}_r\hat{q}_{ir}=\sum_{r=1}^{p}\left(z_{kr}-\bar{z}_r\right)\hat{q}_{ir}.$$

因此

$$\begin{aligned}\operatorname{cov}\left(\boldsymbol{X}_i,\boldsymbol{Z}_j\right)&=\frac{1}{n-1}\sum_{k=1}^{n}\left(x_{ki}-\bar{x}_i\right)\left(z_{kj}-\bar{z}_j\right)\\&=\frac{1}{n-1}\sum_{k=1}^{n}\left\{\left[\sum_{r=1}^{p}\left(z_{kr}-\bar{z}_r\right)\hat{q}_{ir}\right]\left(z_{kj}-\bar{z}_j\right)\right\}\\&=\frac{1}{n-1}\sum_{k=1}^{n}\left\{\sum_{r=1}^{p}\left[\left(z_{kr}-\bar{z}_r\right)\left(z_{kj}-\bar{z}_j\right)\hat{q}_{ir}\right]\right\}\\&=\frac{1}{n-1}\sum_{r=1}^{p}\left\{\sum_{k=1}^{n}\left[\left(z_{kr}-\bar{z}_r\right)\left(z_{kj}-\bar{z}_j\right)\hat{q}_{ir}\right]\right\}\\&=\sum_{r=1}^{p}\left\{\left[\frac{1}{n-1}\sum_{k=1}^{n}\left(z_{kr}-\bar{z}_r\right)\left(z_{kj}-\bar{z}_j\right)\right]\hat{q}_{ir}\right\}\\&=\sum_{r=1}^{p}\operatorname{cov}\left(\boldsymbol{Z}_r,\boldsymbol{Z}_j\right)\hat{q}_{ir}\\&=\operatorname{cov}\left(\boldsymbol{Z}_j,\boldsymbol{Z}_j\right)\hat{q}_{ij}=\hat{q}_{ij}\hat{\lambda}_j.\end{aligned}$$

上式最后两个等式用到了性质(S2)的结果。从而有

$$\begin{aligned}\left(\operatorname{cov}\left(\boldsymbol{X}_i,\boldsymbol{Z}_j\right)\right)_{p\times p}&=\left(\hat{q}_{ij}\hat{\lambda}_j\right)_{p\times p}=\begin{pmatrix}\hat{q}_{11}\hat{\lambda}_1&\hat{q}_{12}\hat{\lambda}_2&\cdots&\hat{q}_{1p}\hat{\lambda}_p\\\hat{q}_{21}\hat{\lambda}_1&\hat{q}_{22}\hat{\lambda}_2&\cdots&\hat{q}_{2p}\hat{\lambda}_p\\\vdots&\vdots&&\vdots\\\hat{q}_{p1}\hat{\lambda}_1&\hat{q}_{p2}\hat{\lambda}_2&\cdots&\hat{q}_{pp}\hat{\lambda}_p\end{pmatrix}\\&=\left(\hat{q}_{ij}\right)_{p\times p}\begin{pmatrix}\hat{\lambda}_1&&&\\&\hat{\lambda}_2&&\\&&\ddots&\\&&&\hat{\lambda}_p\end{pmatrix}=\hat{\boldsymbol{Q}}\hat{\boldsymbol{\Lambda}}.\end{aligned}$$

在证明 $\operatorname{cov}\left(\boldsymbol{X},\boldsymbol{Z}\right)=\hat{\boldsymbol{Q}}\hat{\boldsymbol{\Lambda}}$ 时也可以利用矩阵化的方法来证明。由(A.5)，有

$$\bar{\boldsymbol{X}}=\frac{1}{n}\sum_{k=1}^{n}\boldsymbol{X}_{(k)}=\frac{1}{n}\sum_{k=1}^{n}\hat{\boldsymbol{Q}}\boldsymbol{Z}_{(k)}=\hat{\boldsymbol{Q}}\left(\frac{1}{n}\sum_{k=1}^{n}\boldsymbol{Z}_{(k)}\right)=\hat{\boldsymbol{Q}}\bar{\boldsymbol{Z}},$$

则

$$\boldsymbol{X}_{(k)}-\overline{\boldsymbol{X}}=\hat{\boldsymbol{Q}}\boldsymbol{Z}_{(k)}-\hat{\boldsymbol{Q}}\overline{\boldsymbol{Z}}=\hat{\boldsymbol{Q}}\left(\boldsymbol{Z}_{(k)}-\overline{\boldsymbol{Z}}\right).$$

从而有

$$\begin{aligned}\operatorname{cov}(\boldsymbol{X},\boldsymbol{Z})&=\frac{1}{n-1}\sum_{k=1}^{n}\left(\boldsymbol{X}_{(k)}-\overline{\boldsymbol{X}}\right)\left(\boldsymbol{Z}_{(k)}-\overline{\boldsymbol{Z}}\right)^{\mathrm{T}}\\&=\frac{1}{n-1}\sum_{k=1}^{n}\hat{\boldsymbol{Q}}\left(\boldsymbol{Z}_{(k)}-\overline{\boldsymbol{Z}}\right)\left(\boldsymbol{Z}_{(k)}-\overline{\boldsymbol{Z}}\right)^{\mathrm{T}}\\&=\hat{\boldsymbol{Q}}\left[\frac{1}{n-1}\sum_{k=1}^{n}\left(\boldsymbol{Z}_{(k)}-\overline{\boldsymbol{Z}}\right)\left(\boldsymbol{Z}_{(k)}-\overline{\boldsymbol{Z}}\right)^{\mathrm{T}}\right]\\&=\hat{\boldsymbol{Q}}\operatorname{cov}(\boldsymbol{Z})=\hat{\boldsymbol{Q}}\hat{\boldsymbol{\Lambda}}.\end{aligned}$$

上式最后一个等式用到了性质(S2)的结果。由此发现矩阵化可以大大简化证明。

(S4). 由数据资料阵 $\boldsymbol{X}$ 的样本相关矩阵的定义有

$$\operatorname{cor}(\boldsymbol{X})=\operatorname{cov}\left(\boldsymbol{X}^{*}\right)=\boldsymbol{R}^{X}=\boldsymbol{R}=\left(r_{ij}\right)_{p\times p},$$

由(A.1)，有

$$\operatorname{cov}\left(\boldsymbol{X}^{*}\right)=\operatorname{cov}\left(\boldsymbol{X}^{*},\boldsymbol{X}^{*}\right)=\left(\operatorname{cov}\left(\boldsymbol{X}_{i}^{*},\boldsymbol{X}_{j}^{*}\right)\right)_{p\times p},$$

由(A.3)，有

$$\operatorname{cor}(\boldsymbol{X})=\operatorname{cor}(\boldsymbol{X},\boldsymbol{X})=\left(\operatorname{cor}\left(\boldsymbol{X}_{i},\boldsymbol{X}_{j}\right)\right)_{p\times p}.$$

从而有

$$\operatorname{cor}\left(\boldsymbol{X}_{i},\boldsymbol{X}_{j}\right)=\operatorname{cov}\left(\boldsymbol{X}_{i}^{*},\boldsymbol{X}_{j}^{*}\right)=r_{ij},\quad i,j=1,\cdots,p.$$

(S5). 由矩阵的样本相关矩阵的定义和(A.1)有

$$\operatorname{cor}(\boldsymbol{Z})=\operatorname{cov}\left(\boldsymbol{Z}^{*}\right)=\operatorname{cov}\left(\boldsymbol{Z}^{*},\boldsymbol{Z}^{*}\right)=\left(\operatorname{cov}\left(\boldsymbol{Z}_{i}^{*},\boldsymbol{Z}_{j}^{*}\right)\right)_{p\times p},$$

由(A.3)有

$$\operatorname{cor}(\boldsymbol{Z})=\operatorname{cor}(\boldsymbol{Z},\boldsymbol{Z})=\left(\operatorname{cor}\left(\boldsymbol{Z}_{i},\boldsymbol{Z}_{j}\right)\right)_{p\times p}.$$

由 $\overline{\boldsymbol{Z}^{*}}=0$ 和(S2)的结论有

$$\begin{aligned}\operatorname{cov}\left(\boldsymbol{Z}^{*}\right)&=\frac{1}{n-1}\sum_{k=1}^{n}\left(\boldsymbol{Z}_{(k)}^{*}-\overline{\boldsymbol{Z}^{*}}\right)\left(\boldsymbol{Z}_{(k)}^{*}-\overline{\boldsymbol{Z}^{*}}\right)^{\mathrm{T}}=\frac{1}{n-1}\sum_{k=1}^{n}\boldsymbol{Z}_{(k)}^{*}\boldsymbol{Z}_{(k)}^{*\mathrm{T}}\\&=\frac{1}{n-1}\sum_{k=1}^{n}\hat{\boldsymbol{\Lambda}}^{-\frac{1}{2}}\left(\boldsymbol{Z}_{(k)}-\overline{\boldsymbol{Z}}\right)\left(\boldsymbol{Z}_{(k)}-\overline{\boldsymbol{Z}}\right)^{\mathrm{T}}\hat{\boldsymbol{\Lambda}}^{-\frac{1}{2}}\\&=\hat{\boldsymbol{\Lambda}}^{-\frac{1}{2}}\left[\frac{1}{n-1}\sum_{k=1}^{n}\left(\boldsymbol{Z}_{(k)}-\overline{\boldsymbol{Z}}\right)\left(\boldsymbol{Z}_{(k)}-\overline{\boldsymbol{Z}}\right)^{\mathrm{T}}\right]\hat{\boldsymbol{\Lambda}}^{-\frac{1}{2}}\\&=\hat{\boldsymbol{\Lambda}}^{-\frac{1}{2}}\operatorname{cov}(\boldsymbol{Z})\hat{\boldsymbol{\Lambda}}^{-\frac{1}{2}}=\hat{\boldsymbol{\Lambda}}^{-\frac{1}{2}}\hat{\boldsymbol{\Lambda}}\hat{\boldsymbol{\Lambda}}^{-\frac{1}{2}}=\hat{\boldsymbol{\Lambda}}^{0}=\boldsymbol{E}_{p\times p}=\operatorname{diag}(1,1,\cdots,1)_{p\times p}.\end{aligned}$$

其实我们也可以利用(A.2)和(S2)的结论来证明

$$\begin{aligned}\operatorname{cor}(\boldsymbol{Z})&=\operatorname{cor}(\boldsymbol{Z},\boldsymbol{Z})\\&=\left[\operatorname{Diag}\left(\operatorname{cov}(\boldsymbol{Z})\right)\right]^{-\frac{1}{2}}\operatorname{cov}(\boldsymbol{Z},\boldsymbol{Z})\left[\operatorname{Diag}\left(\operatorname{cov}(\boldsymbol{Z})\right)\right]^{-\frac{1}{2}}\\&=\hat{\Lambda}^{-\frac{1}{2}}\hat{\Lambda}\hat{\Lambda}^{-\frac{1}{2}}=\hat{\Lambda}^{0}=\boldsymbol{E}_{p\times p}.\end{aligned}$$

由此发现式(A.2)在求相关矩阵时非常有力，原因在于它使得我们可以利用已有的结论(S2)来简化证明。式(A.2)在(S6)，(R4)-(R6)的证明中求相关矩阵时都会用到。从而

$$\begin{aligned}&\operatorname{cor}(\boldsymbol{Z}_j)=\operatorname{cov}(\boldsymbol{Z}_j^*)=\operatorname{cov}(\boldsymbol{Z}_j^*,\boldsymbol{Z}_j^*)=1,\quad j=1,\cdots,p,\\&\operatorname{cor}(\boldsymbol{Z}_i,\boldsymbol{Z}_j)=\operatorname{cov}(\boldsymbol{Z}_i^*,\boldsymbol{Z}_j^*)=0,\quad i\neq j,\quad i,j=1,\cdots,p.\end{aligned}$$

(S6). 由(A.3)有

$$\operatorname{cor}(\boldsymbol{X},\boldsymbol{Z})=\left(\operatorname{cor}(\boldsymbol{X}_i,\boldsymbol{Z}_j)\right)_{p\times p}.$$

由矩阵的样本相关矩阵的定义和(A.1)有

$$\operatorname{cor}(\boldsymbol{X},\boldsymbol{Z})=\operatorname{cov}(\boldsymbol{X}^*,\boldsymbol{Z}^*)=\left(\operatorname{cov}(\boldsymbol{X}_i^*,\boldsymbol{Z}_j^*)\right)_{p\times p}.$$

容易验证

$$\begin{aligned}\operatorname{cov}(\boldsymbol{X}^*,\boldsymbol{Z}^*)&=\frac{1}{n-1}\sum_{k=1}^{n}\left(\boldsymbol{X}_{(k)}^*-\overline{\boldsymbol{X}^*}\right)\left(\boldsymbol{Z}_{(k)}^*-\overline{\boldsymbol{Z}^*}\right)^{\mathrm{T}}=\frac{1}{n-1}\sum_{k=1}^{n}\boldsymbol{X}_{(k)}^*\boldsymbol{Z}_{(k)}^{*\mathrm{T}}\\&=\frac{1}{n-1}\sum_{k=1}^{n}\hat{\boldsymbol{D}}^{-\frac{1}{2}}\left(\boldsymbol{X}_{(k)}-\overline{\boldsymbol{X}}\right)\left(\boldsymbol{Z}_{(k)}-\overline{\boldsymbol{Z}}\right)^{\mathrm{T}}\hat{\Lambda}^{-\frac{1}{2}}\\&=\hat{\boldsymbol{D}}^{-\frac{1}{2}}\left[\frac{1}{n-1}\sum_{k=1}^{n}\left(\boldsymbol{X}_{(k)}-\overline{\boldsymbol{X}}\right)\left(\boldsymbol{Z}_{(k)}-\overline{\boldsymbol{Z}}\right)^{\mathrm{T}}\right]\hat{\Lambda}^{-\frac{1}{2}}\\&=\hat{\boldsymbol{D}}^{-\frac{1}{2}}\operatorname{cov}(\boldsymbol{X},\boldsymbol{Z})\hat{\Lambda}^{-\frac{1}{2}}.\end{aligned}$$

由(S3)知

$$\operatorname{cov}(\boldsymbol{X}^*,\boldsymbol{Z}^*)=\hat{\boldsymbol{D}}^{-\frac{1}{2}}\hat{\boldsymbol{Q}}\hat{\Lambda}\hat{\Lambda}^{-\frac{1}{2}}=\hat{\boldsymbol{D}}^{-\frac{1}{2}}\hat{\boldsymbol{Q}}\hat{\Lambda}^{\frac{1}{2}}.$$

其实我们也可以利用(A.2)和(S1)-(S3)的结论来证明

$$\begin{aligned}\operatorname{cor}(\boldsymbol{X},\boldsymbol{Z})&=\left[\operatorname{Diag}\left(\operatorname{cov}(\boldsymbol{X})\right)\right]^{-\frac{1}{2}}\operatorname{cov}(\boldsymbol{X},\boldsymbol{Z})\left[\operatorname{Diag}\left(\operatorname{cov}(\boldsymbol{Z})\right)\right]^{-\frac{1}{2}}\\&=\left[\operatorname{Diag}(\boldsymbol{S})\right]^{-\frac{1}{2}}\hat{\boldsymbol{Q}}\hat{\Lambda}\left[\operatorname{Diag}(\hat{\Lambda})\right]^{-\frac{1}{2}}\\&=\hat{\boldsymbol{D}}^{-\frac{1}{2}}\hat{\boldsymbol{Q}}\hat{\Lambda}\hat{\Lambda}^{-\frac{1}{2}}=\hat{\boldsymbol{D}}^{-\frac{1}{2}}\hat{\boldsymbol{Q}}\hat{\Lambda}^{\frac{1}{2}}.\end{aligned}$$

易知

$$\hat{\boldsymbol{D}}^{-\frac{1}{2}}\hat{\boldsymbol{Q}}\hat{\boldsymbol{\Lambda}}^{\frac{1}{2}}=\begin{pmatrix}\frac{1}{\sqrt{s_{11}}} & & & \\ & \frac{1}{\sqrt{s_{22}}} & & \\ & & \ddots & \\ & & & \frac{1}{\sqrt{s_{pp}}}\end{pmatrix}\left(\hat{q}_{ij}\right)_{p\times p}\begin{pmatrix}\sqrt{\hat{\lambda}_1} & & & \\ & \sqrt{\hat{\lambda}_2} & & \\ & & \ddots & \\ & & & \sqrt{\hat{\lambda}_p}\end{pmatrix}=\left(\hat{q}_{ij}\frac{\sqrt{\hat{\lambda}_j}}{\sqrt{s_{ii}}}\right)_{p\times p}$$

从而有

$$\operatorname{cor}\left(\boldsymbol{X}_i,\boldsymbol{Z}_j\right)=\operatorname{cov}\left(\boldsymbol{X}_i^*,\boldsymbol{Z}_j^*\right)=\hat{q}_{ij}\frac{\sqrt{\hat{\lambda}_j}}{\sqrt{s_{ii}}},\quad i,j=1,\cdots,p.$$

(S7). 由性质(S1)有

$$\sum_{j=1}^{p}\operatorname{cov}\left(\boldsymbol{X}_j\right)=\sum_{j=1}^{p}\operatorname{cov}\left(\boldsymbol{X}_j,\boldsymbol{X}_j\right)=\sum_{j=1}^{p}s_{jj},$$

由性质(S2)有

$$\sum_{j=1}^{p}\operatorname{cov}\left(\boldsymbol{Z}_j\right)=\sum_{j=1}^{p}\operatorname{cov}\left(\boldsymbol{Z}_j,\boldsymbol{Z}_j\right)=\sum_{j=1}^{p}\hat{\lambda}_j,$$

由 $\hat{\boldsymbol{Q}}^{\mathrm{T}}\boldsymbol{S}\hat{\boldsymbol{Q}}=\hat{\boldsymbol{\Lambda}}$ 有

$$\sum_{j=1}^{p}s_{jj}=\operatorname{tr}\left(\boldsymbol{S}\right)=\operatorname{tr}\left(\boldsymbol{S}\hat{\boldsymbol{Q}}\hat{\boldsymbol{Q}}^{\mathrm{T}}\right)=\operatorname{tr}\left(\hat{\boldsymbol{Q}}^{\mathrm{T}}\boldsymbol{S}\hat{\boldsymbol{Q}}\right)=\operatorname{tr}\left(\hat{\boldsymbol{\Lambda}}\right)=\sum_{j=1}^{p}\hat{\lambda}_j.$$

A.3. 从 R 出发求主成分

证明：

(R1). 由(A.1)有

$$\operatorname{cov}\left(\tilde{\boldsymbol{X}}\right)=\operatorname{cov}\left(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{X}}\right)=\left(\operatorname{cov}\left(\tilde{\boldsymbol{X}}_i,\tilde{\boldsymbol{X}}_j\right)\right)_{p\times p},$$

由 $\tilde{\boldsymbol{X}}=\boldsymbol{X}^*$ 及性质(S4)有

$$\operatorname{cov}\left(\tilde{\boldsymbol{X}}\right)=\operatorname{cov}\left(\boldsymbol{X}^*\right)=\operatorname{cor}\left(\boldsymbol{X}\right)=\boldsymbol{R}=\left(r_{ij}\right)_{p\times p},$$

从而有

$$\operatorname{cov}\left(\tilde{\boldsymbol{X}}_i,\tilde{\boldsymbol{X}}_j\right)=\operatorname{cov}\left(\boldsymbol{X}_i^*,\boldsymbol{X}_j^*\right)=r_{ij},\quad i,j=1,\cdots,p.$$

(R2). 由(A.1)有

$$\operatorname{cov}\left(\tilde{\boldsymbol{Z}}\right)=\operatorname{cov}\left(\tilde{\boldsymbol{Z}},\tilde{\boldsymbol{Z}}\right)=\left(\operatorname{cov}\left(\tilde{\boldsymbol{Z}}_i,\tilde{\boldsymbol{Z}}_j\right)\right)_{p\times p}.$$

由于 $\overline{\tilde{\boldsymbol{Z}}}=\boldsymbol{0}$，则

$$
\begin{aligned}
\operatorname{cov}\left(\tilde{\boldsymbol{Z}}\right) &= \frac{1}{n-1}\sum_{k=1}^{n}\left(\tilde{\boldsymbol{Z}}_{(k)}-\bar{\tilde{\boldsymbol{Z}}}\right)\left(\tilde{\boldsymbol{Z}}_{(k)}-\bar{\tilde{\boldsymbol{Z}}}\right)^{\mathrm{T}} = \frac{1}{n-1}\sum_{k=1}^{n}\tilde{\boldsymbol{Z}}_{(k)}\tilde{\boldsymbol{Z}}_{(k)}^{\mathrm{T}} \\
&= \frac{1}{n-1}\sum_{k=1}^{n}\hat{\boldsymbol{Q}}^{*\mathrm{T}}\tilde{\boldsymbol{X}}_{(k)}\tilde{\boldsymbol{X}}_{(k)}^{\mathrm{T}}\hat{\boldsymbol{Q}}^{*} = \hat{\boldsymbol{Q}}^{*\mathrm{T}}\left(\frac{1}{n-1}\sum_{k=1}^{n}\tilde{\boldsymbol{X}}_{(k)}\tilde{\boldsymbol{X}}_{(k)}^{\mathrm{T}}\right)\hat{\boldsymbol{Q}}^{*} \\
&= \hat{\boldsymbol{Q}}^{*\mathrm{T}}\left(\frac{1}{n-1}\sum_{k=1}^{n}\boldsymbol{X}_{(k)}^{*}\boldsymbol{X}_{(k)}^{*\mathrm{T}}\right)\hat{\boldsymbol{Q}}^{*} = \hat{\boldsymbol{Q}}^{*\mathrm{T}}\boldsymbol{R}\hat{\boldsymbol{Q}}^{*} = \hat{\boldsymbol{\Lambda}}^{*},
\end{aligned}
$$

从而

$$
\begin{aligned}
&\operatorname{cov}\left(\tilde{\boldsymbol{Z}}_j\right) = \operatorname{cov}\left(\tilde{\boldsymbol{Z}}_j,\tilde{\boldsymbol{Z}}_j\right) = \hat{\lambda}_j^{*}, \quad j=1,\cdots,p, \\
&\operatorname{cov}\left(\tilde{\boldsymbol{Z}}_i,\tilde{\boldsymbol{Z}}_j\right) = 0, \quad i\neq j, \quad i,j=1,\cdots,p.
\end{aligned}
$$

(R3). 由(A.1)有

$$
\operatorname{cov}\left(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Z}}\right) = \left(\operatorname{cov}\left(\tilde{\boldsymbol{X}}_i,\tilde{\boldsymbol{Z}}_j\right)\right)_{p\times p}.
$$

由于

$$
\bar{\tilde{\boldsymbol{X}}} = \bar{\tilde{\boldsymbol{Z}}} = \boldsymbol{0},
$$

从而

$$
\begin{aligned}
\operatorname{cov}\left(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Z}}\right) &= \frac{1}{n-1}\sum_{k=1}^{n}\left(\tilde{\boldsymbol{X}}_{(k)}-\bar{\tilde{\boldsymbol{X}}}\right)\left(\tilde{\boldsymbol{Z}}_{(k)}-\bar{\tilde{\boldsymbol{Z}}}\right)^{\mathrm{T}} = \frac{1}{n-1}\sum_{k=1}^{n}\tilde{\boldsymbol{X}}_{(k)}\tilde{\boldsymbol{Z}}_{(k)}^{\mathrm{T}} \\
&= \frac{1}{n-1}\sum_{k=1}^{n}\hat{\boldsymbol{Q}}^{*}\tilde{\boldsymbol{Z}}_{(k)}\tilde{\boldsymbol{Z}}_{(k)}^{\mathrm{T}} = \hat{\boldsymbol{Q}}^{*}\left(\frac{1}{n-1}\sum_{k=1}^{n}\tilde{\boldsymbol{Z}}_{(k)}\tilde{\boldsymbol{Z}}_{(k)}^{\mathrm{T}}\right) = \hat{\boldsymbol{Q}}^{*}\operatorname{cov}\left(\tilde{\boldsymbol{Z}}\right),
\end{aligned}
$$

上式的最后一个等式在(R2)中已证明，从而由(R2)的结果有

$$
\operatorname{cov}\left(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Z}}\right) = \hat{\boldsymbol{Q}}^{*}\hat{\boldsymbol{\Lambda}}^{*} = \left(\hat{q}_{ij}^{*}\right)_{p\times p}\begin{pmatrix}\hat{\lambda}_1^{*} & & & \\ & \hat{\lambda}_2^{*} & & \\ & & \ddots & \\ & & & \hat{\lambda}_p^{*}\end{pmatrix} = \left(\hat{q}_{ij}^{*}\hat{\lambda}_j^{*}\right)_{p\times p},
$$

从而

$$
\operatorname{cov}\left(\tilde{\boldsymbol{X}}_i,\tilde{\boldsymbol{Z}}_j\right) = \hat{q}_{ij}^{*}\hat{\lambda}_j^{*}, \quad i,j=1,\cdots,p.
$$

(R4). 由(A.3)有

$$
\operatorname{cor}\left(\tilde{\boldsymbol{X}}\right) = \operatorname{cor}\left(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{X}}\right) = \left(\operatorname{cor}\left(\tilde{\boldsymbol{X}}_i,\tilde{\boldsymbol{X}}_j\right)\right)_{p\times p}.
$$

由 $\tilde{\boldsymbol{X}} = \boldsymbol{X}^{*}$ 及矩阵的样本相关阵的定义有

$$
\operatorname{cor}\left(\tilde{\boldsymbol{X}}\right) = \operatorname{cor}\left(\boldsymbol{X}^{*}\right) = \operatorname{cov}\left(\boldsymbol{X}^{**}\right).
$$

由性质(S4)有

$$
\operatorname{cov}\left(\boldsymbol{X}^{*}\right) = \operatorname{cor}\left(\boldsymbol{X}\right) = \boldsymbol{R} = \left(r_{ij}\right)_{p\times p}.
$$

由(A.4)有

$$
\operatorname{cov}\left(\boldsymbol{X}^{**}\right) = \operatorname{cov}\left(\boldsymbol{X}^{*}\right).
$$

其实我们也可以利用(A.2)和(R1)的结论来证明

$$
\begin{aligned}
\operatorname{cor}(\tilde{\boldsymbol{X}}) &= \operatorname{cor}(\tilde{\boldsymbol{X}}, \tilde{\boldsymbol{X}}) \\
&= \left[\operatorname{Diag}\left(\operatorname{cov}(\tilde{\boldsymbol{X}})\right)\right]^{-\frac{1}{2}} \operatorname{cov}(\tilde{\boldsymbol{X}}, \tilde{\boldsymbol{X}}) \left[\operatorname{Diag}\left(\operatorname{cov}(\tilde{\boldsymbol{X}})\right)\right]^{-\frac{1}{2}} \\
&= \left[\operatorname{Diag}(\boldsymbol{R})\right]^{-\frac{1}{2}} \boldsymbol{R} \left[\operatorname{Diag}(\boldsymbol{R})\right]^{-\frac{1}{2}} = \boldsymbol{E}^{-\frac{1}{2}} \boldsymbol{R} \boldsymbol{E}^{-\frac{1}{2}} = \boldsymbol{R}.
\end{aligned}
$$

从而对 $i, j = 1, \cdots, p,$

$$
\begin{aligned}
\operatorname{cor}(\tilde{\boldsymbol{X}}_i, \tilde{\boldsymbol{X}}_j) &= \operatorname{cor}(\boldsymbol{X}_i^*, \boldsymbol{X}_j^*) = \operatorname{cov}(\boldsymbol{X}_i^{**}, \boldsymbol{X}_j^{**}) = \operatorname{cov}(\boldsymbol{X}_i^*, \boldsymbol{X}_j^*) \\
&= \operatorname{cov}(\tilde{\boldsymbol{X}}_i, \tilde{\boldsymbol{X}}_j) = \operatorname{cor}(\boldsymbol{X}_i, \boldsymbol{X}_j) = r_{ij}.
\end{aligned}
$$

(R5). 由(A.3)有

$$
\operatorname{cor}(\tilde{\boldsymbol{Z}}) = \operatorname{cor}(\tilde{\boldsymbol{Z}}, \tilde{\boldsymbol{Z}}) = \left(\operatorname{cor}(\tilde{\boldsymbol{Z}}_i, \tilde{\boldsymbol{Z}}_j)\right)_{p \times p}.
$$

由矩阵的样本相关阵的定义有

$$
\operatorname{cor}(\tilde{\boldsymbol{Z}}) = \operatorname{cov}(\tilde{\boldsymbol{Z}}^*).
$$

容易验证

$$
\begin{aligned}
&\operatorname{cov}(\tilde{\boldsymbol{Z}}^*) \\
&= \frac{1}{n-1} \sum_{k=1}^{n} \left(\tilde{\boldsymbol{Z}}_{(k)}^* - \overline{\tilde{\boldsymbol{Z}}^*}\right)\left(\tilde{\boldsymbol{Z}}_{(k)}^* - \overline{\tilde{\boldsymbol{Z}}^*}\right)^{\mathrm{T}} = \frac{1}{n-1} \sum_{k=1}^{n} \tilde{\boldsymbol{Z}}_{(k)}^* \tilde{\boldsymbol{Z}}_{(k)}^{*\mathrm{T}} \\
&= \frac{1}{n-1} \sum_{k=1}^{n} (\hat{\Lambda}^*)^{-\frac{1}{2}} \tilde{\boldsymbol{Z}}_{(k)} \tilde{\boldsymbol{Z}}_{(k)}^{\mathrm{T}} (\hat{\Lambda}^*)^{-\frac{1}{2}} = (\hat{\Lambda}^*)^{-\frac{1}{2}} \left(\frac{1}{n-1} \sum_{k=1}^{n} \tilde{\boldsymbol{Z}}_{(k)} \tilde{\boldsymbol{Z}}_{(k)}^{\mathrm{T}}\right) (\hat{\Lambda}^*)^{-\frac{1}{2}} \\
&= (\hat{\Lambda}^*)^{-\frac{1}{2}} \operatorname{cov}(\tilde{\boldsymbol{Z}}) (\hat{\Lambda}^*)^{-\frac{1}{2}}.
\end{aligned}
$$

由性质(R2)的结果有

$$
\operatorname{cov}(\tilde{\boldsymbol{Z}}^*) = (\hat{\Lambda}^*)^{-\frac{1}{2}} \hat{\Lambda}^* (\hat{\Lambda}^*)^{-\frac{1}{2}} = (\hat{\Lambda}^*)^0 = \boldsymbol{E}_{p \times p} = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}_{p \times p}.
$$

其实我们也可以利用(A.2)和(R2)的结论来证明

$$
\begin{aligned}
\operatorname{cor}(\tilde{\boldsymbol{Z}}) &= \operatorname{cor}(\tilde{\boldsymbol{Z}}, \tilde{\boldsymbol{Z}}) \\
&= \left[\operatorname{Diag}\left(\operatorname{cov}(\tilde{\boldsymbol{Z}})\right)\right]^{-\frac{1}{2}} \operatorname{cov}(\tilde{\boldsymbol{Z}}, \tilde{\boldsymbol{Z}}) \left[\operatorname{Diag}\left(\operatorname{cov}(\tilde{\boldsymbol{Z}})\right)\right]^{-\frac{1}{2}} \\
&= (\hat{\Lambda}^*)^{-\frac{1}{2}} \hat{\Lambda}^* (\hat{\Lambda}^*)^{-\frac{1}{2}} = (\hat{\Lambda}^*)^0 = \boldsymbol{E}_{p \times p}.
\end{aligned}
$$

从而

$$
\begin{aligned}
&\operatorname{cor}(\tilde{\boldsymbol{Z}}_j) = \operatorname{cor}(\tilde{\boldsymbol{Z}}_j, \tilde{\boldsymbol{Z}}_j) = 1, \quad j = 1, \cdots, p, \\
&\operatorname{cor}(\tilde{\boldsymbol{Z}}_i, \tilde{\boldsymbol{Z}}_j) = 0, \quad i \neq j, \quad i, j = 1, \cdots, p.
\end{aligned}
$$

(R6). 由(A.3)有

$$\mathrm{cor}\left(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Z}}\right)=\left(\mathrm{cor}\left(\tilde{\boldsymbol{X}}_i,\tilde{\boldsymbol{Z}}_j\right)\right)_{p\times p}.$$

由 $\tilde{\boldsymbol{X}}=\boldsymbol{X}^*$，(A.4)和(A.1)有

$$\mathrm{cov}\left(\tilde{\boldsymbol{X}}^*,\tilde{\boldsymbol{Z}}^*\right)=\mathrm{cov}\left(\boldsymbol{X}^{**},\tilde{\boldsymbol{Z}}^*\right)=\mathrm{cov}\left(\boldsymbol{X}^*,\tilde{\boldsymbol{Z}}^*\right)=\left(\mathrm{cov}\left(\boldsymbol{X}_i^*,\tilde{\boldsymbol{Z}}_j^*\right)\right)_{p\times p}.$$

由矩阵的样本相关阵的定义有

$$\mathrm{cor}\left(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Z}}\right)=\mathrm{cov}\left(\tilde{\boldsymbol{X}}^*,\tilde{\boldsymbol{Z}}^*\right).$$

容易验证

$$\begin{aligned}\mathrm{cov}\left(\tilde{\boldsymbol{X}}^*,\tilde{\boldsymbol{Z}}^*\right)&=\frac{1}{n-1}\sum_{k=1}^{n}\left(\tilde{\boldsymbol{X}}_{(k)}^*-\overline{\tilde{\boldsymbol{X}}^*}\right)\left(\tilde{\boldsymbol{Z}}_{(k)}^*-\overline{\tilde{\boldsymbol{Z}}^*}\right)^{\mathrm{T}}=\frac{1}{n-1}\sum_{k=1}^{n}\tilde{\boldsymbol{X}}_{(k)}^*\tilde{\boldsymbol{Z}}_{(k)}^{*\mathrm{T}}\\&=\frac{1}{n-1}\sum_{k=1}^{n}\tilde{\boldsymbol{X}}_{(k)}\tilde{\boldsymbol{Z}}_{(k)}^{\mathrm{T}}\left(\hat{\Lambda}^*\right)^{-\frac{1}{2}}=\left(\frac{1}{n-1}\sum_{k=1}^{n}\tilde{\boldsymbol{X}}_{(k)}\tilde{\boldsymbol{Z}}_{(k)}^{\mathrm{T}}\right)\left(\hat{\Lambda}^*\right)^{-\frac{1}{2}}\\&=\mathrm{cov}\left(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Z}}\right)\left(\hat{\Lambda}^*\right)^{-\frac{1}{2}}.\end{aligned}$$

再由性质(R3)的结果有

$$\mathrm{cov}\left(\tilde{\boldsymbol{X}}^*,\tilde{\boldsymbol{Z}}^*\right)=\hat{\boldsymbol{Q}}^*\hat{\Lambda}^*\left(\hat{\Lambda}^*\right)^{-\frac{1}{2}}=\hat{\boldsymbol{Q}}^*\left(\hat{\Lambda}^*\right)^{\frac{1}{2}}.$$

其实我们也可以利用(A.2)和(R1)-(R3)的结论来证明

$$\begin{aligned}\mathrm{cor}\left(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Z}}\right)&=\left[\mathrm{Diag}(\mathrm{cov}\left(\tilde{\boldsymbol{X}}\right)\right]^{-\frac{1}{2}}\mathrm{cov}\left(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Z}}\right)\left[\mathrm{Diag}\left(\mathrm{cov}\left(\tilde{\boldsymbol{Z}}\right)\right)\right]^{-\frac{1}{2}}\\&=\left[\mathrm{Diag}\left(\boldsymbol{R}\right)\right]^{-\frac{1}{2}}\hat{\boldsymbol{Q}}^*\hat{\Lambda}^*\left[\mathrm{Diag}\left(\hat{\Lambda}^*\right)\right]^{-\frac{1}{2}}\\&=\boldsymbol{E}^{-\frac{1}{2}}\hat{\boldsymbol{Q}}^*\hat{\Lambda}^*\left(\hat{\Lambda}^*\right)^{-\frac{1}{2}}=\hat{\boldsymbol{Q}}^*\left(\hat{\Lambda}^*\right)^{\frac{1}{2}}.\end{aligned}$$

易知

$$\hat{\boldsymbol{Q}}^*\left(\hat{\Lambda}^*\right)^{\frac{1}{2}}=\left(\hat{q}_{ij}^*\right)_{p\times p}\begin{pmatrix}\sqrt{\hat{\lambda}_1^*}&&&\\&\sqrt{\hat{\lambda}_2^*}&&\\&&\ddots&\\&&&\sqrt{\hat{\lambda}_p^*}\end{pmatrix}=\left(\hat{q}_{ij}^*\sqrt{\hat{\lambda}_j^*}\right)_{p\times p}.$$

从而有

$$\mathrm{cor}\left(\tilde{\boldsymbol{X}}_i,\tilde{\boldsymbol{Z}}_j\right)=\mathrm{cov}\left(\tilde{\boldsymbol{X}}_i^*,\tilde{\boldsymbol{Z}}_j^*\right)=\mathrm{cov}\left(\boldsymbol{X}_i^*,\tilde{\boldsymbol{Z}}_j^*\right)=\hat{q}_{ij}^*\sqrt{\hat{\lambda}_j^*},\quad i,j=1,\cdots,p.$$

(R7). 由(R1)有

$$\sum_{j=1}^{p}\mathrm{cov}\left(\tilde{\boldsymbol{X}}_j\right)=\sum_{j=1}^{p}\mathrm{cov}\left(\tilde{\boldsymbol{X}}_j,\tilde{\boldsymbol{X}}_j\right)=\sum_{j=1}^{p}r_{jj}=\sum_{j=1}^{p}1=p.$$

由(R2)有

$$\sum_{j=1}^{p}\mathrm{cov}\left(\tilde{\boldsymbol{Z}}_j\right)=\sum_{j=1}^{p}\mathrm{cov}\left(\tilde{\boldsymbol{Z}}_j,\tilde{\boldsymbol{Z}}_j\right)=\sum_{j=1}^{p}\hat{\lambda}_j^{*}.$$

由 $\hat{\boldsymbol{Q}}^{*\mathrm{T}}\boldsymbol{R}\hat{\boldsymbol{Q}}^{*}=\hat{\Lambda}^{*}$ 有

$$p=\sum_{j=1}^{p}r_{jj}=\mathrm{tr}\left(\boldsymbol{R}\right)=\mathrm{tr}\left(\boldsymbol{R}\hat{\boldsymbol{Q}}^{*}\hat{\boldsymbol{Q}}^{*\mathrm{T}}\right)=\mathrm{tr}\left(\hat{\boldsymbol{Q}}^{*\mathrm{T}}\boldsymbol{R}\hat{\boldsymbol{Q}}^{*}\right)=\mathrm{tr}\left(\hat{\Lambda}^{*}\right)=\sum_{j=1}^{p}\hat{\lambda}_j^{*}.$$

A.4. 数值模拟

在这里给出例 1 和例 2 的详细的数值模拟结果。

下面我们举两个例子来验证(S1)-(S7)和(R1)-(R7)的正确性。

例 1.

$$\boldsymbol{X}=\left(x_{ij}\right)_{n\times p},\ n=4,\ p=3,\ x_{ij}\sim N\left(0,1\right).$$

为重复本文的结果，我们在 R 软件中使用 set.seed(1)，此时

$$\boldsymbol{X}=\begin{pmatrix}-0.6264538 & 0.3295078 & 0.5757814\\ 0.1836433 & -0.8204684 & -0.3053884\\ -0.8356286 & 0.4874291? & 1.5117812\\ 1.5952808 & 0.7383247 & 0.3898432\end{pmatrix}$$

$$\boldsymbol{S}=\begin{pmatrix}1.2147560 & 0.1184091 & -0.4100692\\ 0.1184091 & 0.4764913 & 0.3553357\\ -0.4100692 & 0.3553357 & 0.5609438\end{pmatrix},$$

$$\boldsymbol{R}=\begin{pmatrix}1.0000000 & 0.1556371 & -0.4967672\\ 0.1556371 & 1.0000000 & 0.6873084\\ -0.4967672 & 0.6873084 & 1.0000000\end{pmatrix},$$

$$\hat{\Lambda}=\begin{pmatrix}\hat{\lambda}_1 & & \\ & \hat{\lambda}_2 & \\ & & \hat{\lambda}_3\end{pmatrix}=\begin{pmatrix}1.41507491 & & \\ & 0.79365619 & \\ & & 0.04346005\end{pmatrix},$$

$$\hat{\boldsymbol{Q}}=\left(\hat{\boldsymbol{a}}_1,\hat{\boldsymbol{a}}_2,\hat{\boldsymbol{a}}_3\right)=\begin{pmatrix}0.89017519 & 0.3354400 & 0.3083312\\ -0.05874925 & 0.7555882 & -0.6524070\\ -0.45181485 & 0.5626423 & 0.6923128\end{pmatrix}.$$

由于 $\boldsymbol{S}$ 的特征值全为正，故 $\boldsymbol{S}>\boldsymbol{0}$。

(S1).

$$\mathrm{cov}\left(\boldsymbol{X}\right)=\boldsymbol{S}=\begin{pmatrix}1.2147560 & 0.1184091 & -0.4100692\\ 0.1184091 & 0.4764913 & 0.3553357\\ -0.4100692 & 0.3553357 & 0.5609438\end{pmatrix},$$

(S2).

$$\mathrm{cov}\left(\boldsymbol{Z}\right)=\hat{\Lambda}=\begin{pmatrix}1.41507491 & & \\ & 0.79365619 & \\ & & 0.04346005\end{pmatrix},$$

(S3).

$$\operatorname{cov}(\boldsymbol{X},\boldsymbol{Z})=\hat{\boldsymbol{Q}}\hat{\Lambda}=\begin{pmatrix}1.25966458 & 0.2662241 & 0.01340009\\ -0.08313458 & 0.5996773 & -0.02835364\\ -0.63935186 & 0.4465446 & 0.03008795\end{pmatrix},$$

(S4).

$$\operatorname{cor}(\boldsymbol{X})=\boldsymbol{R}=\begin{pmatrix}1.0000000 & 0.1556371 & -0.4967672\\ 0.1556371 & 1.0000000 & 0.6873084\\ -0.4967672 & 0.6873084 & 1.0000000\end{pmatrix},$$

(S5).

$$\operatorname{cor}(\boldsymbol{Z})=\boldsymbol{E}_{3\times 3}=\begin{pmatrix}1 & & \\ & 1 & \\ & & 1\end{pmatrix},$$

(S6).

$$\operatorname{cor}(\boldsymbol{X},\boldsymbol{Z})=\hat{\boldsymbol{D}}^{-\frac{1}{2}}\hat{\boldsymbol{Q}}\hat{\Lambda}^{\frac{1}{2}}=\left(\operatorname{Diag}(\boldsymbol{S})\right)^{-\frac{1}{2}}\hat{\boldsymbol{Q}}\hat{\Lambda}^{\frac{1}{2}}$$
$$=\begin{pmatrix}0.9607727 & 0.2711356 & 0.05832007\\ -0.1012429 & 0.9751556 & -0.19703176\\ -0.7176134 & 0.6692508 & 0.19270268\end{pmatrix},$$

(S7).

$$\operatorname{tr}\left(\operatorname{cov}(\boldsymbol{X})\right)=\operatorname{tr}\left(\operatorname{cov}(\boldsymbol{Z})\right)=2.252191.$$

此时

$$\tilde{\boldsymbol{X}}=\boldsymbol{X}^{*}=\begin{pmatrix}-0.64025588 & 0.2112313 & 0.04376326\\ 0.09475296 & -1.4547163 & -1.13275832\\ -0.83004217 & 0.4400087 & 1.29349294\\ 1.37554509 & 0.8034763 & -0.20449788\end{pmatrix},$$

$$\hat{\Lambda}^{*}=\begin{pmatrix}\hat{\lambda}_1^{*} & & \\ & \hat{\lambda}_2^{*} & \\ & & \hat{\lambda}_3^{*}\end{pmatrix}=\begin{pmatrix}1.77908842 & & \\ & 1.14726052 & \\ & & 0.07365107\end{pmatrix},$$

$$\hat{\boldsymbol{Q}}^{*}=\left(\hat{\boldsymbol{a}}_1^{*},\hat{\boldsymbol{a}}_2^{*},\hat{\boldsymbol{a}}_3^{*}\right)=\begin{pmatrix}-0.3534130 & 0.81512585 & 0.4589870\\ 0.5781696 & 0.57604641 & -0.5778326\\ 0.7354041 & -0.06115876 & 0.6748633\end{pmatrix}.$$

由于 $\boldsymbol{R}$ 的特征值全为正，故 $\boldsymbol{R}>\boldsymbol{0}$。

(R1).

$$\operatorname{cov}(\tilde{\boldsymbol{X}})=\boldsymbol{R}=\begin{pmatrix}1.0000000 & 0.1556371 & -0.4967672\\ 0.1556371 & 1.0000000 & 0.6873084\\ -0.4967672 & 0.6873084 & 1.0000000\end{pmatrix},$$

(R2).

$$\operatorname{cov}(\tilde{Z})=\hat{\Lambda}^{*}=\begin{pmatrix}1.77908842 & & \\ & 1.14726052 & \\ & & 0.07365107\end{pmatrix},$$

(R3).

$$\operatorname{cov}(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Z}})=\hat{\boldsymbol{Q}}^{*}\hat{\Lambda}^{*}=\begin{pmatrix}-0.628753 & 0.93516170 & 0.03380488\\ 1.028615 & 0.66087530 & -0.04255799\\ 1.308349 & -0.07016503 & 0.04970440\end{pmatrix},$$

(R4).

$$\operatorname{cor}(\tilde{\boldsymbol{X}})=\boldsymbol{R}=\begin{pmatrix}1.0000000 & 0.1556371 & -0.4967672\\ 0.1556371 & 1.0000000 & 0.6873084\\ -0.4967672 & 0.6873084 & 1.0000000\end{pmatrix},$$

(R5).

$$\operatorname{cor}(\tilde{\boldsymbol{Z}})=\boldsymbol{E}_{3\times 3}=\begin{pmatrix}1 & & \\ & 1 & \\ & & 1\end{pmatrix},$$

(R6).

$$\operatorname{cor}(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Z}})=\hat{\boldsymbol{Q}}^{*}\left(\hat{\Lambda}^{*}\right)^{\frac{1}{2}}=\begin{pmatrix}-0.4713911 & 0.8730833 & 0.1245632\\ 0.7711769 & 0.6170047 & -0.1568164\\ 0.9809002 & -0.0655073 & 0.1831493\end{pmatrix},$$

(R7).

$$\operatorname{tr}\left(\operatorname{cov}(\tilde{\boldsymbol{X}})\right)=3=\operatorname{tr}\left(\operatorname{cov}(\tilde{\boldsymbol{Z}})\right).$$

由数值结果可以看出，(S1)-(S7)和(R1)-(R7)均是正确的。

下面给一个 $\boldsymbol{S}\geq\boldsymbol{0}$ 且 $\boldsymbol{R}\geq\boldsymbol{0}$ 的例子，即 $\boldsymbol{S}$ 和 $\boldsymbol{R}$ 都有 1 个 0 特征值，此时(S5)，(S6)，(R5)和(R6)是不成立的，但其余性质是成立的。

例 2.

$$\boldsymbol{X}=\left(x_{ij}\right)_{n\times p},\ n=p=3,\ x_{ij}\sim N(0,1).$$

为重复本文的结果，我们在 R 软件中使用 set.seed(1)，此时

$$\boldsymbol{X}=\begin{pmatrix}-0.6264538 & 1.5952808 & 0.4874291\\ 0.1836433 & 0.3295078 & 0.7383247\\ -0.8356286 & -0.8204684 & 0.5757814\end{pmatrix}$$

$$\boldsymbol{S}=\begin{pmatrix}0.28982112 & 0.10867553 & 0.05840744\\ 0.10867553 & 1.46007844 & -0.05734883\\ 0.05840744 & -0.05734883 & 0.01619585\end{pmatrix},$$

$$\boldsymbol{R}=\begin{pmatrix}1.0000000 & 0.1670624 & 0.8525136\\ 0.1670624 & 1.0000000 & -0.3729362\\ 0.8525136 & -0.3729362 & 1.0000000\end{pmatrix},$$

$$\hat{\boldsymbol{\Lambda}}=\begin{pmatrix}\hat{\lambda}_1 & & \\ & \hat{\lambda}_2 & \\ & & \hat{\lambda}_3\end{pmatrix}=\begin{pmatrix}1.471929 & & \\ & 0.2941663 & \\ & & -7.139463e-17\end{pmatrix},$$

$$\hat{\boldsymbol{Q}}=\left(\hat{\boldsymbol{a}}_1,\hat{\boldsymbol{a}}_2,\hat{\boldsymbol{a}}_3\right)=\begin{pmatrix}-0.08974467 & 0.97207413 & -0.21683584\\ -0.99532799 & -0.07975163 & 0.05442315\\ 0.03561033 & 0.22070697 & 0.97468987\end{pmatrix}.$$

由于 $\hat{\lambda}_3=0$，故 $\boldsymbol{S}\geq\boldsymbol{0}$。

(S1).

$$\operatorname{cov}(\boldsymbol{X})=\boldsymbol{S}=\begin{pmatrix}0.28982112 & 0.10867553 & 0.05840744\\ 0.10867553 & 1.46007844 & -0.05734883\\ 0.05840744 & -0.05734883 & 0.01619585\end{pmatrix},$$

(S2).

$$\operatorname{cov}(\boldsymbol{Z})=\hat{\boldsymbol{\Lambda}}=\begin{pmatrix}1.471929 & & \\ & 0.2941663 & \\ & & 0\end{pmatrix},$$

(S3).

$$\operatorname{cov}(\boldsymbol{X},\boldsymbol{Z})=\hat{\boldsymbol{Q}}\hat{\boldsymbol{\Lambda}}=\begin{pmatrix}-0.13209779 & 0.28595150 & 0\\ -1.46505219 & -0.02346024 & 0\\ 0.05241588 & 0.06492456 & 0\end{pmatrix},$$

(S4).

$$\operatorname{cor}(\boldsymbol{X})=\boldsymbol{R}=\begin{pmatrix}1.0000000 & 0.1670624 & 0.8525136\\ 0.1670624 & 1.0000000 & -0.3729362\\ 0.8525136 & -0.3729362 & 1.0000000\end{pmatrix},$$

(S5).

$$\operatorname{cor}(\boldsymbol{Z})=\begin{pmatrix}1 & 0 & -0.7037615\\ 0 & 1 & -0.4139885\\ -0.7037615 & -0.4139885 & 1\end{pmatrix}\neq\begin{pmatrix}1 & & \\ & 1 & \\ & & 1\end{pmatrix}=\boldsymbol{E}_{3\times 3},$$

(S6).

$$\operatorname{cor}(\boldsymbol{X},\boldsymbol{Z})=\begin{pmatrix}-0.2022493 & 0.97933406 & -0.2630977\\ -0.9993591 & -0.03579711 & 0.7181300\\ 0.3394826 & 0.94061233 & -0.6283174\end{pmatrix}$$
$$\neq\begin{pmatrix}-0.2022493 & 0.97933406 & \text{NaN}\\ -0.9993591 & -0.03579711 & \text{NaN}\\ 0.3394826 & 0.94061233 & \text{NaN}\end{pmatrix}$$
$$=\hat{\boldsymbol{D}}^{-\frac{1}{2}}\hat{\boldsymbol{Q}}\hat{\boldsymbol{\Lambda}}^{\frac{1}{2}}=\left(\operatorname{Diag}(\boldsymbol{S})\right)^{-\frac{1}{2}}\hat{\boldsymbol{Q}}\hat{\boldsymbol{\Lambda}}^{\frac{1}{2}},$$

(S7).

$$\operatorname{tr}\left(\operatorname{cov}(\boldsymbol{X})\right)=\operatorname{tr}\left(\operatorname{cov}(\boldsymbol{Z})\right)=1.766095.$$

此时

$$\tilde{\boldsymbol{X}}=\boldsymbol{X}^{*}=\begin{pmatrix}-0.3720764 & 1.01558921 & -0.8885750\\ 1.1327004 & -0.03194387 & 1.0828999\\ -0.7606241 & -0.98364534 & -0.1943249\end{pmatrix},$$

$$\hat{\boldsymbol{\Lambda}}^{*}=\begin{pmatrix}\hat{\lambda}_1^{*} & & \\ & \hat{\lambda}_2^{*} & \\ & & \hat{\lambda}_3^{*}\end{pmatrix}=\begin{pmatrix}1.879171 & & \\ & 1.120829 & \\ & & 0\end{pmatrix},$$

$$\hat{\boldsymbol{Q}}^{*}=\left(\hat{\boldsymbol{a}}_1^{*},\hat{\boldsymbol{a}}_2^{*},\hat{\boldsymbol{a}}_3^{*}\right)=\begin{pmatrix}0.6669312 & -0.3826931 & 0.6393346\\ -0.1799815 & -0.9153613 & -0.3601670\\ 0.7230556 & 0.1251382 & -0.6793607\end{pmatrix}.$$

由于 $\hat{\lambda}_3^{*}=0$，故 $\boldsymbol{R}\geq\boldsymbol{0}$。

(R1).

$$\operatorname{cov}(\tilde{\boldsymbol{X}})=\boldsymbol{R}=\begin{pmatrix}1.0000000 & 0.1670624 & 0.8525136\\ 0.1670624 & 1.0000000 & -0.3729362\\ 0.8525136 & -0.3729362 & 1.0000000\end{pmatrix},$$

(R2).

$$\operatorname{cov}(\tilde{\boldsymbol{Z}})=\hat{\boldsymbol{\Lambda}}^{*}=\begin{pmatrix}1.879171 & & \\ & 1.120829 & \\ & & 0\end{pmatrix},$$

(R3).

$$\operatorname{cov}(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Z}})=\hat{\boldsymbol{Q}}^{*}\hat{\boldsymbol{\Lambda}}^{*}=\begin{pmatrix}1.253278 & -0.4289336 & 0\\ -0.338216 & -1.0259635 & 0\\ 1.358745 & 0.1402585 & 0\end{pmatrix},$$

(R4).

$$\operatorname{cor}(\tilde{\boldsymbol{X}})=\boldsymbol{R}=\begin{pmatrix}1.0000000 & 0.1670624 & 0.8525136\\ 0.1670624 & 1.0000000 & -0.3729362\\ 0.8525136 & -0.3729362 & 1.0000000\end{pmatrix},$$

(R5).

$$\operatorname{cor}\left(\tilde{\boldsymbol{Z}}\right)=\begin{pmatrix}1 & 0 & -0.4178971\\ 0 & 1 & -0.9084944\\ -0.4178971 & -0.9084944 & 1\end{pmatrix}\neq\begin{pmatrix}1 & & \\ & 1 & \\ & & 1\end{pmatrix}=\boldsymbol{E}_{3\times 3},$$

(R6).

$$\operatorname{cor}\left(\tilde{\boldsymbol{X}},\tilde{\boldsymbol{Z}}\right)=\begin{pmatrix}0.9142484 & -0.4051542 & -0.01398142\\ -0.2467238 & -0.9690858 & 0.98351419\\ 0.9911853 & 0.1324828 & -0.53457336\end{pmatrix}$$

$$\neq\begin{pmatrix}0.9142484 & -0.4051542 & 0\\ -0.2467238 & -0.9690858 & 0\\ 0.9911853 & 0.1324828 & 0\end{pmatrix}=\hat{\boldsymbol{Q}}^{*}\left(\hat{\Lambda}^{*}\right)^{\frac{1}{2}},$$

(R7).

$$\operatorname{tr}\left(\operatorname{cov}\left(\tilde{\boldsymbol{X}}\right)\right)=3=\operatorname{tr}\left(\operatorname{cov}\left(\tilde{\boldsymbol{Z}}\right)\right).$$

由数值结果可以看出，除了(S5)，(S6)，(R5)和(R6)之外，(S1)-(S4)，(S7)和(R1)-(R4)，(R7)均是正确的。