1.6 HW 6

作业 6 链接

练习 1.1

假设数据矩阵为 $X_{n\times p} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)^\top$，且 X 已经中心化。假设样本协方差矩阵 $S = X^\top X/(n-1)$ 的谱分解为 $S = V\Lambda V^\top$，其中 $V^\top V = VV^\top = I_p$，$\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_p)$，$\lambda_1 \geq \cdots \geq \lambda_p \geq 0$。主成分矩阵 $Y_{n\times p} = XV$ 的第 j 列为所有 n 个样本点的第 j 主成分，$j = 1, \ldots, p$。

(a) 证明 Y 是中心化的矩阵，其样本协方差矩阵为 $Y^\top Y/(n-1) = \Lambda$。Y 与 X 的样本协方差矩阵为 $Y^\top X/(n-1) = \Lambda V^\top$（即第 k 个变量与第 j 个主成分之间的样本协方差为 $v_{kj}\lambda_j$，其中 v_{kj} 是 V 的第 (k,j) 元素）。

(b) 证明第 j 主成分和第 k 个变量的样本相关系数 $r_{jk} = v_{kj}\sqrt{\lambda_j/s_{kk}}$，其中 s_{kk} 为第 k 个变量的样本方差，即 S 的第 (k,k) 元素。

解

由于 X 已经中心化，故 $X^\top \mathbf{1}_n = 0$，所以

$$Y^\top \mathbf{1}_n = V^\top X^\top \mathbf{1}_n = 0$$

即 Y 也是列中心化的矩阵。

(a) 计算样本协方差矩阵：

$$\frac{1}{n-1}Y^\top Y = \frac{1}{n-1}V^\top X^\top XV = V^\top SV = V^\top(V\Lambda V^\top)V = (V^\top V)\Lambda(V^\top V) = \Lambda.$$

再计算 $Y^\top X$ 的样本协方差矩阵：

$$\frac{1}{n-1}Y^\top X = \frac{1}{n-1}V^\top X^\top X = V^\top S = V^\top V\Lambda V^\top = \Lambda V^\top.$$

因此，第 j 主成分与第 k 个变量之间的样本协方差为 $\left(\Lambda V^\top\right)_{jk} = \lambda_j v_{kj}$。

(b) 样本相关系数定义为：

$$r_{jk} = \frac{\mathrm{Cov}(Y_j, X_k)}{\sqrt{\mathrm{Var}(Y_j)}\sqrt{\mathrm{Var}(X_k)}} = \frac{\lambda_j v_{kj}}{\sqrt{\lambda_j}\sqrt{s_{kk}}} = v_{kj}\sqrt{\frac{\lambda_j}{s_{kk}}}.$$

练习 1.2

假设 $\mathbf{x}_1, \ldots, \mathbf{x}_n \in \mathbb{R}^p$ 的样本方差-协方差矩阵具有如下形式：

$$S = \sigma^2 I_p + \mathbf{1}_p\mathbf{1}_p^\top + \mathbf{e}\mathbf{e}^\top,$$

其中 $\mathbf{1}_p = (1, 1, \ldots, 1)^\top$，$\mathbf{e} = (1, -1, 0, \ldots, 0)^\top$。求 $\mathbf{x}_i, i = 1, \ldots, n$ 的第一主成分和第二主成分，以及它们的方差在总方差中所占的累计比例。

解

$$\begin{aligned}
S &= \sigma^2 I_p + \mathbf{1}_p\mathbf{1}_p^\top + \mathbf{e}\mathbf{e}^\top \\
S\mathbf{1}_p &= \sigma^2\mathbf{1}_p + p\mathbf{1}_p = (\sigma^2 + p)\mathbf{1}_p \\
S\mathbf{e} &= \sigma^2\mathbf{e} + 2\mathbf{e} = (\sigma^2 + 2)\mathbf{e}
\end{aligned}$$

对于任意满足 $\mathbf{v} \perp \mathbf{1}$, $\mathbf{v} \perp \mathbf{e}$ 的向量 $\mathbf{v} \in \mathcal{C}(\mathbf{1}, \mathbf{e})^\perp$（即 $\mathbf{1}$ 与 $\mathbf{e}$ 张成空间的正交补空间），都有：

$$S\mathbf{v} = \sigma^2\mathbf{v}$$

因为 $\dim(\mathcal{C}(\mathbf{1}, \mathbf{e})^\perp) = p - 2$，所以 σ^2 是特征值，重数为 $p-2$，所以 S 的所有特征值为

$$\sigma^2 + p, \quad \sigma^2 + 2, \quad \underbrace{\sigma^2, \ldots, \sigma^2}_{p-2\text{ 个}}$$

第一主成分

$$y_1 = \frac{1}{\sqrt{p}}\mathbf{1}_p^\top(x_1, \ldots, x_n) = \frac{1}{\sqrt{p}}\mathbf{1}_p^\top X^\top$$

第二主成分

$$y_2 = \frac{1}{\sqrt{2}} \mathbf{e}^\top (x_1, \dots, x_n) = \frac{1}{\sqrt{2}} \mathbf{e}^\top X^\top$$

前两个主成分在总方差中所占的累计比例为

$$\frac{\sigma^2 + p + \sigma^2 + 2}{\sigma^2 + p + \sigma^2 + 2 + (p-2)\sigma^2} = \frac{2\sigma^2 + p + 2}{p\sigma^2 + p + 2}$$

练习 1.3 Thurstone（1933）关于 PMA（primary mental abilities，基本心理能力）的研究中，对 $n = 213$ 名被试进行了 9 项与语言表达能力有关的测试。9 项测试分别是：

Sentences, Vocabulary, Sentence Completion, First Letters, Four Letter Words, Suffixes, Letter Series, Pedigrees, Letter Group

9 项测试的样本相关系数矩阵如下表所示：

	Sentences	*Vocab*	*Sent.Comp*	*First.Letters*	*Four.Letters*	*Suffixes*	*Letter.Series*	*Pedigrees*	*Letter.Group*
Sentences	1.000	0.828	0.776	0.439	0.432	0.447	0.447	0.541	0.380
Vocab	0.828	1.000	0.779	0.493	0.464	0.489	0.432	0.537	0.358
Sent.Comp	0.776	0.779	1.000	0.460	0.425	0.443	0.401	0.534	0.359
First.Letters	0.439	0.493	0.460	1.000	0.674	0.590	0.381	0.350	0.424
Four.Letters	0.432	0.464	0.425	0.674	1.000	0.541	0.402	0.367	0.446
Suffixes	0.447	0.489	0.443	0.590	0.541	1.000	0.288	0.320	0.325
Letter.Series	0.447	0.432	0.401	0.381	0.402	0.288	1.000	0.555	0.598
Pedigrees	0.541	0.537	0.534	0.350	0.367	0.320	0.555	1.000	0.452
Letter.Group	0.380	0.358	0.359	0.424	0.446	0.325	0.598	0.452	1.000

上述 9 项测试任务的目的是希望分别测量语言理解能力（*Verbal comprehension*, 项目 1–3）、词语流利度（*Word fluency*, 项目 4–6）和推理能力（*Reasoning*, 项目 7–9）。

由于没有原始数据，只有相关系数矩阵，我们对该矩阵进行主成分分析（PCA）。以下是 R 命令 'princomp(covmat = Thurstone)' 的输出结果：

```
Call:
princomp(covmat = Thurstone)

Standard deviations:
Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7 Comp.8 Comp.9
 2.202  1.044  1.019  0.690  0.669  0.612  0.567  0.484  0.409
```

前 3 个主成分的载荷为：

	Sentences	Vocab	Sent.Comp	First.Letters	Four.Letter	Suffixes	Letter.Series	Pedigrees	Letter.Group
Comp.1	0.37	0.38	0.36	0.33	0.32	0.30	0.30	0.32	0.29
Comp.2	0.38	0.33	0.36	-0.44	-0.45	-0.33	-0.04	0.23	-0.24
Comp.3	0.15	0.21	0.19	0.21	0.13	0.36	-0.57	-0.33	-0.52

请回答以下问题：

(a) 试计算前 3 个主成分的方差贡献率。

(b) 试根据成分载荷解释前三个主成分的含义，并判断是否与 Thurstone 的目标一致？

解

(a) 计算前三个主成分的方差贡献率：

根据题目中的输出结果，主成分的标准差如下：

$$\text{Comp.1} = 2.202,$$
$$\text{Comp.2} = 1.044,$$
$$\text{Comp.3} = 1.019.$$

因此，每个主成分对应的方差为标准差的平方。由于使用的是相关系数矩阵（每个变量的方差为 1），总方差为 9。因此，前三个主成分的方差贡献率分别为：

$$
\begin{aligned}
\text{贡献率}_1 &= \frac{2.202^2}{9} \approx 0.5388, \\
\text{贡献率}_2 &= \frac{1.044^2}{9} \approx 0.1211, \\
\text{贡献率}_3 &= \frac{1.019^2}{9} \approx 0.1154.
\end{aligned}
$$

累计贡献率为：

$$0.5388 + 0.1211 + 0.1154 = 0.7753.$$

(b) 从载荷表分析：

- 主成分 **1**：所有变量载荷均为正，且前三项（Verbal）和中间三项（Fluency）略高，说明该主成分代表一种综合语言能力，结合了理解与流利度。
- 主成分 **2**：前三个变量（Verbal）载荷为正，第四至第六个变量（Fluency）载荷为负，说明此主成分区分语言理解与词语流利度，是一种对立维度。
- 主成分 **3**：第七至第九个变量（Reasoning）载荷为负且绝对值较大，说明该主成分主要衡量推理能力。

因此，前三个主成分可以解释为：

- Comp.1：语言能力总水平（Verbal + Fluency）
- Comp.2：Verbal 与 Fluency 的区分因子
- Comp.3：推理能力因子

这与 Thurstone 的理论目标一致，他希望识别出三种心理能力：语言理解、流利度与推理。因此我们可以认为：主成分分析得到的结果与 Thurstone 的目标基本一致。

练习 1.4 我们应用主成分分析方法研究欧洲 23 个国家的首都城市的气温数据集 `temperature`。23 个首都的信息如下表所示：

首都城市	国家	位置	首都城市	国家	位置	首都城市	国家	位置
Amsterdam (阿姆斯特丹)	荷兰	W	Athens (雅典)	希腊	S	Berlin (柏林)	德国	W
Brussels (布鲁塞尔)	比利时	W	Budapest (布达佩斯)	匈牙利	E	Copenhagen (哥本哈根)	丹麦	N
Dublin (都柏林)	爱尔兰	N	Helsinki (赫尔辛基)	芬兰	N	Kiev (基辅)	乌克兰	E
Krakow (克拉科夫)	波兰旧都	E	Lisbon (里斯本)	葡萄牙	S	London (伦敦)	英国	N
Madrid (马德里)	西班牙	S	Minsk (明斯克)	白俄罗斯	E	Moscow (莫斯科)	俄罗斯	E
Oslo (奥斯陆)	挪威	N	Paris (巴黎)	法国	W	Prague (布拉格)	捷克	E
Reykjavik (雷克雅未克)	冰岛	N	Rome (罗马)	意大利	S	Sarajevo (萨拉热窝)	波黑	S
Sofia (索菲亚)	保加利亚	E	Stockholm (斯德哥尔摩)	瑞典	N			

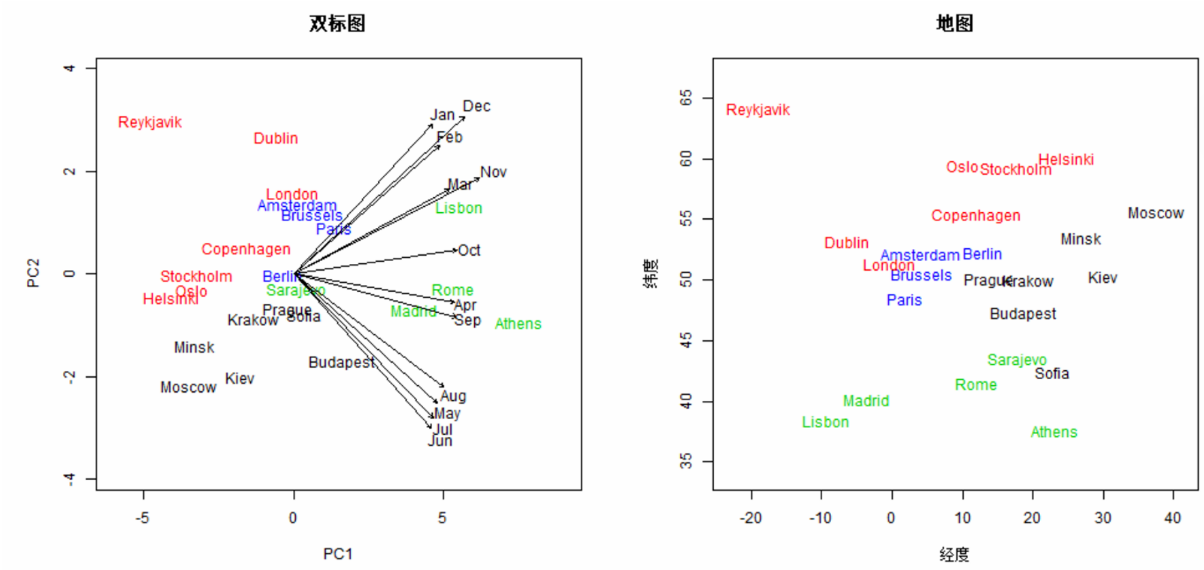
其中地理位置 W, E, S, N 分别表示欧洲西、东、南、北部，每个城市记录了历史上 12 个月份（Jan–Dec）的平均气温。对标准化后的数据应用主成分分析方法（R: `princomp`），部分输出结果即主成分的标准差（Standard deviations）如下所示（方框内）：

```
Call: princomp(x = temperature, cor = T)
Standard deviations:
Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7
  3.16   1.36   0.36   0.20   0.13   0.11   0.08
Comp.8 Comp.9 Comp.10 Comp.11 Comp.12
  0.05   0.03    0.03    0.02    0.01
```

前两个主成分的载荷如下：

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Comp.1	0.27	0.28	0.30	0.31	0.28	0.26	0.27	0.29	0.31	0.31	0.30	0.28
Comp.2	0.39	0.34	0.21	-0.07	-0.34	-0.40	-0.37	-0.30	-0.11	0.06	0.21	0.35

左下图是主成分分析的双标图，右下图给出了这 23 个城市的经纬度坐标图。



(a) 试计算第一、二主成分在总方差中的占比。根据载荷数据解释第一、二主成分的含义，它们与地理位置有什么关系？

(b) 四月份和哪个月份的气温相关系数最大？哪个或哪几个月份的温度最能反映南北差异？

(c) 哪个城市或地区春秋两季（四月和九月）的温度相对较高？萨拉热窝（Sarajevo）是南欧城市，其气温与其它南欧城市（地图中的绿色标记城市）相差较远，该城市与其它南欧城市在地理环境上有什么差异？

解

(a) 计算第一、二主成分在总方差中的占比，并解释含义与地理位置的关系：

第一、二主成分的贡献率为：

$$\frac{3.16^2}{12} \approx 0.832, \quad \frac{1.36^2}{12} \approx 0.154$$

累计解释比例为：

$$0.832 + 0.154 = 0.986$$

即前两主成分解释了大约 98.6% 的总方差。

根据主成分载荷可知：

第一个主成分对所有月份的载荷均为正且相近（$0.26 \sim 0.31$），表示其衡量的是全年平均温度的总体水平。因此，它主要反映的是南北差异。

第二个主成分在春秋为负，在冬夏为正，表示它衡量的是季节温差的大小，反映的是大陆性和海洋性气候差异。

因此，第一主成分与南北纬度高度相关，第二主成分则与海陆距离（东西）更相关。

(b) 由载荷信息可知，2 月和 9 月的温度变化最相似（相关性最大），因为它们在主成分空间中投影方向接近。此外，4 月、9 月、10 月这几个月在第一主成分方向上接近平行，说明它们在该方向（代表南北差异）上具有类似的结构。这表明这些月份的气温在南北城市之间差异较大，能较好地反映南北纬度的影响。

(c) 从数据中可以看出，**Athens**、**Rome**、**Madrid** 在 4 月和 9 月的温度相对较高，表现出典型的南欧地中海型气候特征，春秋温暖、冬季较湿润。

相比之下，**Sarajevo** 虽然地理上位于南欧，但在主成分图中与南欧其他城市差异明显。其在 4 月和 9 月的温度较低，靠近中欧或内陆城市。

这可能是因为 **Sarajevo** 地处高海拔的内陆盆地，气候更偏向大陆性气候，距离地中海较远，海洋调节作用弱。因此该城市在春秋季与典型南欧城市相比，温差较大，昼夜变化更显著。

练习 1.5 法国 Decastar 巡回赛是世界上最大的国际田联十项全能（decathlon）赛事，比赛次序如下：

第一天：100 米、跳远、铅球、跳高、400 米；

第二天：110 米栏、铁饼、撑杆跳、标枪、1500 米。

其中径赛跑步类以时间量度成绩（单位：秒，时间值越小越好），田赛包括跳跃和投掷项目，以高度或距离量度成绩（单位：米，值越大越好）。对 2004 年该赛事前 13 名运动员 10 项成绩的相关系数矩阵做主成分分析，前两个主成分 PC1 和 PC2 的载荷（即主成分方向）如下所示：

	100 米	跳远	铅球	跳高	400 米	110 米栏	铁饼	撑杆跳	标枪	1500 米
PC1	0.48	-0.37	-0.20	-0.19	0.34	0.50	-0.31	0.03	-0.01	-0.31
PC2	0.03	0.00	-0.43	0.27	-0.33	-0.20	-0.19	-0.53	-0.33	-0.40

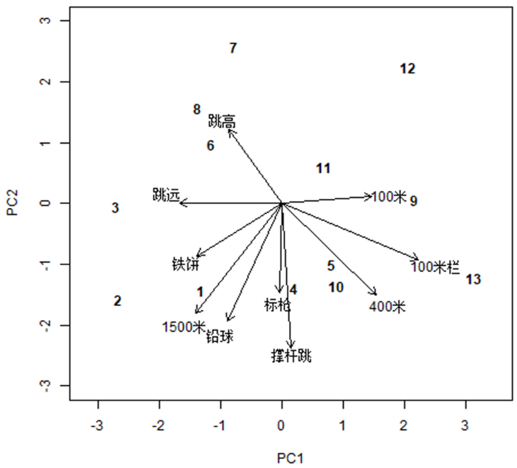
请回答下列问题：

(a) 根据上述载荷，解释第一主成分 PC1 的含义。

(b) 已知前两个主成分 PC1 和 PC2 的标准差分别为 1.76, 1.42，计算它们的累计方差贡献率。

(c) 分析运动员的 PC 散点图（右图，数字代表运动员，大小代表名次），第一名（数字 1）的投掷类成绩如何（好、中、差）？他的 4 个径赛项目表现各如何（好、中、差）？最后一名（数字 13）有什么特点？

(d) 跳高和撑杆跳是正相关还是负相关？图中跳远和 100 米方向相反说明了什么？1500 米作为径赛项目与其它径赛项目类似吗？简单解释原因。



解

(a) 观察载荷表可见：短跑项目的载荷为正（数值越大代表表现越差），大部分田赛项目载荷为负（数值越大代表表现越好）。而长跑耐力项目的载荷为负（数值越大代表表现越差）。因此，**PC1** 实质衡量了爆发力/耐力：

(b)

$$\text{累计贡献率} = \frac{1.76^2 + 1.42^2}{10} \approx 51.14\%.$$

(c) 运动员 1 投掷项目好，径赛项目除了 1500 m 表现较差之外其他都较好。
运动员 13 只有撑杆跳，标枪和 1500 m 表现较好，其余项目表现都较差。

(d) (i) 跳高 *vs.* 撑杆跳：箭头几乎相反 → 负相关。
(ii) 跳远 *vs. 100 m*：两箭头方向相反。由于 100 m 以 “时间” 计，箭头正向表示 “慢”。因而箭头相反恰好体现短跑越快 (时间越小)，往往跳远越远。这说明爆发力项目之间的共通性。
(iii) *1500 m* 与其他径赛：1500 m 与其余径赛没有太大的相似性，更需要耐力。

练习 1.6 假设随机向量 $\mathbf{x} = (x_1, x_2, x_3)^\top$ 的协方差矩阵（相关系数矩阵）为：

$$\Sigma = \begin{pmatrix} 1 & 0.63 & 0.45 \\ 0.63 & 1 & 0.35 \\ 0.45 & 0.35 & 1 \end{pmatrix}$$

假设 $\mathbf{x} = (x_1, x_2, x_3)^\top$ 由如下单因子模型产生：

$$\begin{aligned} x_1 &= 0.9F + \varepsilon_1 \\ x_2 &= 0.7F + \varepsilon_2 \\ x_3 &= 0.5F + \varepsilon_3 \end{aligned}$$

其中 $\mathrm{var}(F) = 1, \quad \mathrm{cov}(F, \boldsymbol{\varepsilon}) = \mathbf{0}$。

请写出载荷矩阵 L，并求 $\Psi = \mathrm{var}(\boldsymbol{\varepsilon})$ 的估计。

解

$$L = (0.9, 0.7, 0.5)^\top, \quad \Psi = \Sigma - LL^T = \begin{pmatrix} 0.19 & 0 & 0 \\ 0 & 0.51 & 0 \\ 0 & 0 & 0.75 \end{pmatrix}$$

练习 1.7

假设 $\mathbf{x} = (x_1, \ldots, x_6)^\top$ 的相关系数矩阵如下：

$$R = \begin{pmatrix} 1.000 & & & & & \\ 0.505 & 1.000 & & & & \\ 0.569 & 0.422 & 1.000 & & & \\ 0.602 & 0.467 & 0.926 & 1.000 & & \\ 0.621 & 0.482 & 0.877 & 0.874 & 1.000 & \\ 0.603 & 0.450 & 0.878 & 0.894 & 0.937 & 1.000 \end{pmatrix}$$

对两因子模型，使用极大似然法得到如下因子载荷矩阵：

变量	F_1	F_2
x_1	0.478	0.417
x_2	0.371	0.323
x_3	0.593	0.727
x_4	0.514	0.855
x_5	0.859	0.506
x_6	0.735	0.604

请试计算：

(a) 每个变量的特殊方差。
(b) 每个变量的公共方差。
(c) 每个因子解释的总方差比例。
(d) 每个变量被 F_1, F_2 解释的方差比例。
(e) 残差矩阵：$R - LL^\top - \Psi$。

解

(a) 特殊方差 $\Psi_i = \sigma_{ii}^2 - (l_{i1}^2 + \cdots + l_{im}^2), i = 1, \cdots, 6$ 代入数据为 $(0.598, 0.759, 0.120, 0.005, 0.006, 0.095)$
(b) 公共方差 $h_i = l_{i1}^2 + \cdots + l_{im}^2, i = 1, \cdots, 6$ 代入数据为 $(0.402, 0.241, 0.880, 0.995, 0.994, 0.905)$
(c) 记 L 的各列：$L = (\mathbf{l}_{(1)}, \mathbf{l}_{(2)})$，$SS_j = \left\|\mathbf{l}_{(j)}\right\|^2 = l_{1j}^2 + \cdots + l_{pj}^2$ 为 L 第 j 列的平方和，第 j 个因子的方差贡献率率为 $SS_j/\mathrm{tr}(R), j = 1, 2$，代入数据得: $(37.63\%, 35.98\%)$
(d) 第 i 个变量被 F_j 解释的方差的比例记为 $\beta_{ij} := \frac{h_{ij}^2}{\sigma_{ii}^2}$ 代入数据得

$$\beta_{i1} = (0.228, 0.137, 0.351, 0.264, 0.737, 0.540)$$

$$\beta_{i2} = (0.174, 0.104, 0.528, 0.731, 0.256, 0.365)$$

(e)

$$R - LL^T - \Psi = \begin{pmatrix} 0 & 0.192971 & -0.01761 & -0.00023 & -0.0006 & -0.0002 \\ 0.192971 & 0 & -0.03282 & 0.000141 & -0.00013 & -0.01778 \\ -0.01761 & -0.03282 & 0 & -0.00039 & -0.00025 & 0.003037 \\ -0.00023 & 0.000141 & -0.00039 & 0 & -0.00016 & -0.00021 \\ -0.0006 & -0.00013 & -0.00025 & -0.00016 & 0 & 0.000011 \\ -0.0002 & -0.01778 & 0.003037 & -0.00021 & 0.000011 & 0 \end{pmatrix}$$

练习 1.8 对第 3 题 Thurstone 数据进行 3 因子分析，得到如下载荷（空白处为 0 或接近 0）:

```
Call:
factanal(factors = 3, covmat = Thurstone, rotation = "promax")

Loadings:
                   Factor1 Factor2 Factor3
Sentences           0.919
Vocabulary          0.901
Sent.Completion     0.842
First.Letters               0.876
Four.Letter.Words           0.756
Suffixes            0.174   0.645   -0.103
Letter.Series                       0.893
Pedigrees           0.347   0.494
Letter.Group                0.184    0.675
```

请回答下列问题:
(a) 三个因子是否与第 3 题的主成分的含义相同或接近?
(b) 计算三个因子的累计方差贡献率，它一定小于主成分方法的贡献率，解释原因。
(c) 分别计算前 3 个、中间 3 个和最后 3 个变量的 3 因子方差解释比例。

练习 1.9

(a) 三个因子与第 3 题主成分含义接近:

- 因子 1 主要载于 Sentences, Vocabulary, Sent.Completion，与主成分 1 高度一致。
- 因子 2 主要载于 First.Letters, Four.Letter.Words，与主成分 2 结构相符（符号方向可能相反）。
- 因子 3 主要载于 Letter.Series, Pedigrees, Letter.Group，与主成分 3 意义相近。

(b)

$$
\begin{aligned}
SS_1 &= 0.919^2 + 0.901^2 + 0.842^2 + 0.174^2 + 0.347^2 + (-0.116)^2 = 2.52947,\\
SS_2 &= 0.876^2 + 0.756^2 + 0.645^2 + 0.184^2 = 1.78879,\\
SS_3 &= 0.893^2 + 0.494^2 + (-0.103)^2 + 0.675^2 = 1.50772.
\end{aligned}
$$

因子贡献率为

$$\frac{SS_1}{9} \approx 0.2811, \quad \frac{SS_2}{9} \approx 0.1988, \quad \frac{SS_3}{9} \approx 0.1675,$$

累计贡献率

$$0.2811 + 0.1988 + 0.1675 = 0.6474$$

因子分析仅针对各变量的公共方差（剔除了特有方差与误差项）进行提取，而主成分分析则以最大化全部原始方差（包括公共方差和特有方差）为目标构造主成分，因此其累计方差贡献率必然大于或等于因子分析法。

(c)

$$
\begin{aligned}
&\text{前 3 个：}\quad \text{平均} = \frac{0.919^2 + 0.901^2 + 0.842^2}{3} = 0.788.\\
&\text{中 3 个：}\quad \text{平均} = \frac{0.876^2 + 0.756^2 + 0.174^2 + 0.645^2 + (-0.103)^2}{3} = 0.599.\\
&\text{后 3 个：}\quad \text{平均} = \frac{0.893^2 + 0.347^2 + 0.494^2 + (-0.116)^2 + 0.184^2 + 0.675^2}{3} = 0.555.
\end{aligned}
$$