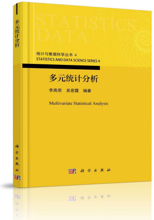
第10章
因子分析

李高荣

北京师范大学统计学院

E-mail: *ligaorong@bnu.edu.cn*



1. 因子分析模型
2. 因子载荷矩阵的估计方法
   - 主成分法
   - 主因子法
   - 极大似然法
3. 因子旋转
4. 因子分析模型的拟合优度检验
5. 因子得分
   - Thomson因子得分
   - Bartlett因子得分
   - Thomson因子得分和Bartlett因子得分比较
   - 案例与R语言计算
6. 因子分析与主成分分析的关系

- 扫二维码获取在线课件和相关教学电子资源
- 请遵守电子资源使用协议

因子分析(factor analysis)的思想：

因子分析是一种数据简化或降维的技术，它通过研究众多变量之间的内部依赖关系，探求观测数据中的基本结构，并用少数几个假想变量来表示其基本的数据结构。

- 因子分析在心理学、社会学、经济学等学科都取得成功的应用。

- 设$X_1, \cdots, X_p$为某个多元总体下的p个可观测的连续型随机变量
- $F_1, \cdots, F_k$ 为潜在因子，这里因子的个数$k < p$
- 每个可观测变量X_i可被表达为：

$$X_i - \mu_i = a_{i1}F_1 + \cdots + a_{ik}F_k + \varepsilon_i, \quad i = 1, \cdots, p,$$

  - ▷ μ_i为X_i的期望
  - ▷ $F_1, \cdots, F_k$ 刻画的是所有可观测变量$X_1, \cdots, X_p$ 潜在的共同特征，故被称为公共因子
  - ▷ ε_i刻画的是X_i不可被公共因子解释的特有部分，故称为X_i的特殊因子
  - ▷ 系数a_{ij}反应了第j个公共因子F_j对第i个变量X_i的重要程度，称其为载荷

- 假设公共因子$F_1, \cdots, F_k$两两不相关，且$E(F_i)=0$ 和$\mathrm{Var}(F_i)=1$
- 假设特殊因子$\varepsilon_1, \cdots, \varepsilon_p$两两不相关，且$E(\varepsilon_i)=0$ 和$\mathrm{Var}(\varepsilon_i)=\psi_i$
- 公共因子和特殊因子相互独立。首先引入一些记号，令

$$\boldsymbol{X}=(X_1, \cdots, X_p)', \qquad \boldsymbol{\mu}=(\mu_1, \cdots, \mu_p)',$$
$$\boldsymbol{F}=(F_1, \cdots, F_k)', \qquad \boldsymbol{\varepsilon}=(\varepsilon_1, \cdots, \varepsilon_p)',$$

$$\mathbf{A}=\begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{p1} & \cdots & a_{pk} \end{pmatrix}, \qquad \boldsymbol{\Psi}=\begin{pmatrix} \psi_1 & & \\ & \ddots & \\ & & \psi_p \end{pmatrix}.$$

定义10.1：正交因子分析模型

设$\boldsymbol{X}$为p维随机向量，$\boldsymbol{\mu}$为其均值向量。若$\boldsymbol{X}$ 可随机表示为：

$$\boldsymbol{X} = \mathbf{A}\boldsymbol{F} + \boldsymbol{\mu} + \boldsymbol{\varepsilon}, \tag{1.1}$$

其中$\mathbf{A}$是一个$p \times k$的常数矩阵($k < p$)，$\boldsymbol{F}$和$\boldsymbol{\varepsilon}$分别为k维和p维随机向量，且

$$E(\boldsymbol{F}) = \mathbf{0}, \qquad \mathrm{Cov}(\boldsymbol{F}) = \mathbf{I}_k,$$

$$E(\boldsymbol{\varepsilon}) = \mathbf{0}, \qquad \mathrm{Cov}(\boldsymbol{\varepsilon}) = \boldsymbol{\Psi},$$

$$\mathrm{Cov}(\boldsymbol{F}, \boldsymbol{\varepsilon}) = \mathbf{0},$$

则称由式(1.1)定义的模型称为$\boldsymbol{X}$ 的正交因子分析模型，称$\mathbf{A}$ 为因子载荷矩阵，称$\boldsymbol{F}$为$\boldsymbol{X}$的公共因子向量，称$\boldsymbol{\varepsilon}$为$\boldsymbol{X}$的特殊因子向量。

- 在正交因子分析模型下，$\boldsymbol{\Sigma} = (\sigma_{ij})$ 可表示为：

$$\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}' + \boldsymbol{\Psi}.$$

※ 载荷矩阵的统计意义：注意到

$$\begin{aligned}\mathrm{Cov}(\boldsymbol{X}, \boldsymbol{F}) &= E\Big[\boldsymbol{X} - E(\boldsymbol{X})\Big]\Big[\boldsymbol{F} - E(\boldsymbol{F})\Big]' = E\Big[(\boldsymbol{X} - \boldsymbol{\mu})\boldsymbol{F}'\Big] \\ &= E\Big[(\mathbf{A}\boldsymbol{F} + \boldsymbol{\varepsilon})\boldsymbol{F}'\Big] = \mathbf{A}E(\boldsymbol{F}\boldsymbol{F}') + E(\boldsymbol{\varepsilon}\boldsymbol{F}') = \mathbf{A}.\end{aligned}$$

▷ 因此，有$\mathrm{Cov}(X_i, F_j) = a_{ij}$，其中$i = 1, \cdots, p$; $j = 1, \cdots, k$

▷ 每个载荷a_{ij} 就是$\boldsymbol{X}$的第i个分量X_i与公共因子F_j的协方差，反映了第i个变量与第j 个公共因子的相关重要性

※ 变量共同度的统计意义：易见

$$\mathrm{Var}(X_i) = \sigma_{ii} = \sum_{j=1}^{k} a_{ij}^2 + \psi_i =: h_i + \psi_i.$$

- ▷ h_i 称为共同度，是由与其他变量共享的公共因子所引起的X_i的方差部分，反映了X_i对公共因子的依赖程度
- ▷ ψ_i 是只与X_i有关的特殊因子的方差
- ▷ 如果h_i非常靠近σ_{ii}，则ψ_i 非常小，表明因子分析的效果好，从原变量空间到公共因子空间的转化性质好

※ 公共因子方差贡献的统计意义：因子载荷矩阵$\mathbf{A}$中各列元素的平方和

$$\sum_{i=1}^{p} a_{ij}^2$$

称为公共因子F_j对X的各个分量$X_1, \cdots, X_p$的总的方差贡献，是衡量公共因子F_j的相对重要性的依据。

- 因子分析的本质是对X的协方差阵结构的一种简化，一般适用于因子个数k小于维数p的情形
- 当$k < (p-1)/2$时，因子分析可将协方差阵从$p(p+1)/2$维参数空间降维到$p(k+1)$维参数空间

- 例： 在实际应用中，常采用刻画组内相关结构的协方差矩阵

$$\boldsymbol{\Sigma} = \sigma^2 \begin{pmatrix} 1 & \cdots & \rho \\ \vdots & \ddots & \vdots \\ \rho & \cdots & 1 \end{pmatrix}_{p\times p}.$$

- 考虑一种简单的因子分析模型，即$k=1$。事实上，$\boldsymbol{\Sigma}$可以写成：

$$\boldsymbol{\Sigma} = (\sigma\sqrt{\rho}\mathbf{1}_p)(\sigma\sqrt{\rho}\mathbf{1}_p)' + \sigma^2(1-\rho)\mathbf{I}_p.$$

- 故$\boldsymbol{X}$可用单因子分析模型来表达，即

$$\boldsymbol{X} - \boldsymbol{\mu} = \sigma\sqrt{\rho}\mathbf{1}_p F + \boldsymbol{\varepsilon}$$

- F为公共因子，$\boldsymbol{\varepsilon}$为特殊因子向量，且$\mathrm{Cov}(\boldsymbol{\varepsilon}) = \sigma^2(1-\rho)\mathbf{I}_p$

性质10.1.1

因子载荷矩阵不唯一。

证明：如果因子分析模型成立，则将因子$\boldsymbol{F}$正交旋转后，即对任意$k\times k$正交阵$\mathbf{Q}$，都有：$\boldsymbol{X}=(\mathbf{AQ})(\mathbf{Q}'\boldsymbol{F})+\boldsymbol{\mu}+\boldsymbol{\varepsilon}$. 记$\mathbf{A}^*=\mathbf{AQ}$和$\boldsymbol{F}^*=\mathbf{Q}'\boldsymbol{F}$。易验证：

$$\boldsymbol{\Sigma}=\mathbf{A}^*(\mathbf{A}^*)'+\boldsymbol{\Psi},$$

$$E(\boldsymbol{F}^*)=E(\mathbf{Q}'\boldsymbol{F})=\mathbf{0},\qquad \mathrm{Cov}(\boldsymbol{F}^*)=\mathbf{Q}'\mathbf{Q}=\mathbf{I}_k,$$

$$\mathrm{Cov}(\boldsymbol{F}^*,\boldsymbol{\varepsilon})=\mathbf{Q}'\mathrm{Cov}(\boldsymbol{F},\boldsymbol{\varepsilon})=\mathbf{0}.$$

因此，$\boldsymbol{X}=\mathbf{A}^*\boldsymbol{F}^*+\boldsymbol{\mu}+\boldsymbol{\varepsilon}$ 也是$\boldsymbol{X}$的因子结构模型，其中$\mathbf{A}^*$为因子载荷矩阵，$\boldsymbol{F}^*$为公共因子。表明因子载荷矩阵$\mathbf{A}$并不是唯一的。

性质10.1.2

因子分析具有尺度不变性。

证明：记$\mathbf{C} = \mathrm{diag}(c_1, \cdots, c_p)$，其中$c_i$为随机向量$\boldsymbol{X}$第$i$个分量$X_i$发生尺度变换的比例。于是随机向量$\boldsymbol{X}$在新尺度下可表示为$\boldsymbol{Y} = \mathbf{C}\boldsymbol{X}$。

因此，如果因子分析模型成立，则在新尺度下向量$\boldsymbol{Y}$的因子分析模型为：

$$\mathbf{C}\boldsymbol{X} - \mathbf{C}\boldsymbol{\mu} = \mathbf{C}\mathbf{A}\boldsymbol{F} + \mathbf{C}\boldsymbol{\varepsilon},$$

$$\mathrm{Cov}(\boldsymbol{Y}) = \mathbf{C}\boldsymbol{\Sigma}\mathbf{C}' = \mathbf{C}\mathbf{A}\mathbf{A}'\mathbf{C}' + \mathbf{C}\boldsymbol{\Psi}\mathbf{C}',$$

其中$\mathbf{CA}$为因子载荷矩阵，$\boldsymbol{F}$为公共因子，且

$$\mathbf{C\Psi C}' = \begin{pmatrix} c_1^2\psi_1 & & \\ & \ddots & \\ & & c_p^2\psi_p \end{pmatrix} =: \mathbf{\Psi}_{\boldsymbol{Y}}$$

为特殊因子$\mathbf{C}\varepsilon$的协方差矩阵。

显然$\boldsymbol{Y}$的载荷矩阵$\mathbf{CA}$就是原向量$\boldsymbol{X}$的因子载荷矩阵$\mathbf{A}$重新尺度化，即将$\mathbf{A}$的第i行乘以c_i，同样对于特殊因子的方差$\mathbf{\Psi}_{\boldsymbol{Y}}$也可通过对原向量$\boldsymbol{X}$的特殊因子的方差$\mathbf{\Psi}$重新尺度化得到，即$\mathbf{\Psi}$的第$i$个对角元乘以$c_i$的平方，公共因子$\boldsymbol{F}$不变。

- 在实际应用中，随机向量$\boldsymbol{X}$的各分量的量纲往往不一致，因此，往往需要将变量标准化
- 通过相关系数矩阵$\mathbf{R}$求因子分析模型，即令

$$\mathbf{C}=\begin{pmatrix} \frac{1}{\sqrt{\sigma_{11}}} & & \\ & \ddots & \\ & & \frac{1}{\sqrt{\sigma_{pp}}} \end{pmatrix}.$$

- 于是，标准化后的向量$\boldsymbol{Y}$的协方差矩阵为$\mathbf{C}\boldsymbol{\Sigma}\mathbf{C}'=\mathbf{R}$，这里，$\sigma_{11},\cdots,\sigma_{pp}$为$\boldsymbol{\Sigma}$的主对角元素

- 因子分析就是来确定因子分析模型中未知的因子载荷矩阵$\mathbf{A}$、公共因子$\boldsymbol{F}$ 以及特殊因子的方差矩阵$\boldsymbol{\Psi}$
- 重点介绍三种求解因子载荷矩阵$\mathbf{A}$的三种方法：
  1. 主成分法
  2. 主因子法
  3. 极大似然法

- 设随机向量$\boldsymbol{X}$的协方差阵为$\boldsymbol{\Sigma}$，$\lambda_1 \geq \cdots \geq \lambda_p$ 为$\boldsymbol{\Sigma}$的特征值，$\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p$ 为对应的标准化特征向量，则

$$\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}' + \boldsymbol{\Psi} = \mathbf{U}\boldsymbol{\Lambda}\mathbf{U}' = \sum_{i=1}^{p} \lambda_i \boldsymbol{u}_i \boldsymbol{u}_i',$$

这里，$\mathbf{U} = (\boldsymbol{u}_1, \cdots, \boldsymbol{u}_p)$，$\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \cdots, \lambda_p)$。

- 若前k个特征值$\lambda_1, \cdots, \lambda_k$的累计贡献率很高，令

$$\mathbf{U}_1 = (\boldsymbol{u}_1, \cdots, \boldsymbol{u}_k), \qquad \mathbf{U}_2 = (\boldsymbol{u}_{k+1}, \cdots, \boldsymbol{u}_p),$$

$$\boldsymbol{\Lambda}_1 = \mathrm{diag}(\lambda_1, \cdots, \lambda_k), \qquad \boldsymbol{\Lambda}_2 = \mathrm{diag}(\lambda_{k+1}, \cdots, \lambda_p).$$

- 于是：$\boldsymbol{\Sigma} = \mathbf{U}_1\boldsymbol{\Lambda}_1\mathbf{U}_1' + \mathbf{U}_2\boldsymbol{\Lambda}_2\mathbf{U}_2'$.
- 由于$\mathbf{U}_1\boldsymbol{\Lambda}_1\mathbf{U}_1'$ 解释了$\boldsymbol{X}$的主要相关关系，故因子载荷矩阵$\mathbf{A}$的估计为：

$$\widehat{\mathbf{A}} = \mathbf{U}_1\boldsymbol{\Lambda}_1^{1/2}.$$

- 特殊因子的方差ψ_i可用$\mathbf{U}_2\boldsymbol{\Lambda}_2\mathbf{U}_2'$的第$i$个对角元素来估计，等价于

$$\widehat{\psi}_i = \sigma_{ii} - \widehat{h}_i,$$

其中$\widehat{h}_i$ 为X_i的共同度h_i的估计，即为$\mathbf{U}_1\boldsymbol{\Lambda}_1\mathbf{U}_1'$的第$i$个对角线元素，把该估计方法称为主成分法。

- 注意到：若$\lambda_{k+1} = \cdots = \lambda_p = 0$，则
$$\boldsymbol{\Sigma} = \mathbf{U}_1 \boldsymbol{\Lambda}_1 \mathbf{U}_1' = \mathbf{A}\mathbf{A}'.$$
- 通过主成分法得到的载荷矩阵的估计$\widehat{\mathbf{A}} = \mathbf{U}_1 \boldsymbol{\Lambda}_1^{1/2}$ 就是真实的载荷矩阵$\mathbf{A}$。
- 因此，当$\lambda_{k+1}, \cdots, \lambda_p$相对很小时，主成分法可得到较好的载荷矩阵的估计。
- 如果$\boldsymbol{\Sigma}$未知，在上述估计中，可用$\boldsymbol{\Sigma}$的无偏估计$\mathbf{S}$ 替代即可。
- 为了消除量纲，需要对变量进行标准化，这时对$\widehat{\mathbf{R}}$求解主成分即可。

- 例. 美国债券指数的因子分析模型：考虑期限为30年、20年、10年、5年和1年的美国债券指数的月对数收益率。
- 数据来自GRSP数据库1942年1月至1999年12月的月数据，共有696个月的数据。
- 由于数据具有序列依赖性，所以本例基于消除数据序列依赖性后的相关系数矩阵

$$
\mathbf{R} = \begin{pmatrix} 1.0 & & & & \\ 0.98 & 1.0 & & & \\ 0.92 & 0.92 & 1.0 & & \\ 0.85 & 0.86 & 0.90 & 1.0 & \\ 0.66 & 0.67 & 0.71 & 0.84 & 1.0 \end{pmatrix}.
$$

- 从相关系数矩阵可以看出5个指标之间的两两相关系数都比较大，所以可能存在公共因子。
- 设这5个不同期限的美国债券指数的对数收益率为$X_1, \cdots, X_5$。
- 取$k=2$，考虑因子分析模型：

$$X_i - E(X_i) = a_{i1}F_1 + a_{i2}F_2 + \varepsilon_i, \qquad i = 1, 2, \cdots, 5.$$

- 应用主成分法求该模型下的因子载荷矩阵和公共因子方差等指标。
- 解：根据主成分法，可编写R语言函数fa.pcm()，实现因子载荷矩阵的主成分法估计。

```
fa.pcm = function(S, k){
  p = nrow(S); diag_S = diag(S); sum_rank = sum(diag_S)
  rowname = rownames(S); colname = paste('Factor',1:k,sep='')
  A = matrix(0,nrow=p,ncol=k,dimnames=list(rowname,colname))
  eig = eigen(S)
  for (i in 1:k)
  A[,i] = sqrt(eig$values[i])*eig$vectors[,i]
  h = diag(A%*%t(A))
  rowname = c('SS loadings','Proportion Var','Cumulative Var')
  B = matrix(0,nrow=3,ncol=k,dimnames=list(rowname,colname))
  for (i in 1:k){
    B[1,i] = sum(A[,i]^2); B[2,i] = B[1,i]/sum_rank
    B[3,i] = sum(B[1,1:i])/sum_rank
 }
  A = round(A,3); B = round(B,3); h = round(h,3)
  diag_S = round(diag_S,3); method = c('Principal Component Method')
  list(method=method, loadings=A, var=cbind(common=h, spcific=diag_S-h), B=B)
}
```

```
names = c('30年', '20年', '10年', '5年', '1年')
R   = matrix(c(1.0, 0.98, 0.92, 0.85, 0.66, 0.98,
              1.0, 0.92, 0.86, 0.67, 0.92, 0.92,
              1.0, 0.90, 0.71, 0.85, 0.86, 0.90,
              1.0, 0.84, 0.66, 0.67, 0.71, 0.84, 1),
              nrow = 5, ncol = 5,
              dimnames=list(names, names), byrow=T)
fa.pcm.fit = fa.pcm(R, k=2)
#### 输出结果:
$method
[1] "Principal Component Method"
$loadings
     Factor1 Factor2
30年  -0.952   0.258
20年  -0.956   0.242
10年  -0.960   0.129
5年   -0.956  -0.156
1年   -0.825  -0.547
```

```
$var
      common  spcific
30年   0.973    0.027
20年   0.973    0.027
10年   0.937    0.063
5年    0.938    0.062
1年    0.980    0.020
$B
                Factor1 Factor2
SS loadings       4.337   0.465
Proportion Var    0.867   0.093
Cumulative Var    0.867   0.960
```

- 由上分析结果可知：利用主成分法估得的因子载荷矩阵为：

$$\mathbf{A}=\begin{pmatrix} -0.952 & -0.956 & -0.960 & -0.956 & -0.825 \\ 0.258 & 0.242 & 0.129 & -0.156 & -0.547 \end{pmatrix}'$$

- 各变量的共同度都很高，分别为0.973、0.973、0.937、0.938 和0.980，说明5个变量的各自方差都可由这两个因子很好地解释。
- 两个因子的累计贡献率达到了96%，反映了用该因子模型进行分析是合理的。

- 所有5种债券收益对第一个因子的载荷都接近于-1，因此，第一个因子可归结为“债券收益率"，体现了一般的美国债券收益率；
- 对第二个因子的载荷与期限长短有关的，且长期和短期的符号相反，因子载荷的和接近于0，故第二个因子可归结为“期限"效应，可解释为长期债券与短期债券的比较。

- 由于因子分析具有尺度不变性，为了简单，我们将仅考虑相关系数矩阵$\mathbf{R}$的因子载荷矩阵的估计问题，即考虑$\boldsymbol{X}$标准化后的因子结构：

$$\mathbf{R} = \mathbf{A}\mathbf{A}' + \mathbf{\Psi}$$

⊛ 目的： 利用主因子法讨论$\mathbf{A}$和$\mathbf{\Psi}$ 的估计问题

- 如果X相关系数矩阵$\mathbf{R}$和所有特殊因子方差ψ_i都已知，则有

$$\mathbf{R}-\mathbf{\Psi}=\mathbf{A}\mathbf{A}',$$

称$\mathbf{R}^*=\mathbf{R}-\mathbf{\Psi}$为约相关矩阵。

- 与相关系数矩阵相比，仅在主对角元素有差异，$\mathbf{R}^*$的主对角元素是变量的共同度h_i，而不是1，且$\mathbf{R}^*$仍为半正定矩阵。
- 主因子法就是对约相关矩阵$\mathbf{R}^*$进行主成分法。

- 首先对$\mathbf{R}^*$ 进行谱分解，得

$$\mathbf{R}^* = \mathbf{QDQ}' = \sum_{i=1}^{p} d_i \boldsymbol{q}_i \boldsymbol{q}_i',$$

这里，$\mathbf{D} = \mathrm{diag}(d_1, \cdots, d_p)$，记$d_1 \geq \cdots \geq d_p$。

- 类似主成分法，只用$\mathbf{R}^*$的分解式中前k个较大的特征值对应项的和来估计$\mathbf{A}$，即

$$\mathbf{R}^* = \mathbf{AA}' \approx \sum_{i=1}^{k} d_i \boldsymbol{q}_i \boldsymbol{q}_i' = \mathbf{Q}_1 \mathbf{D}_1 \mathbf{Q}_1'.$$

- 则，$\mathbf{A}$的估计为
$$\widehat{\mathbf{A}} = \mathbf{Q}_1\mathbf{D}_1^{1/2},$$
其中$\mathbf{Q}_1 = (\boldsymbol{q}_1, \cdots, \boldsymbol{q}_k)$和$\mathbf{D}_1 = \mathrm{diag}(d_1, \cdots, d_k)$。
- 问题：在实际应用中，X的相关系数矩阵$\mathbf{R}$和特殊因子方差矩阵$\mathbf{\Psi}$都是未知的。
- 解决办法：给定共同度h_i的初值$\widehat{h}_i$，然后用初值$\widehat{h}_i$替换样本相关阵$\widehat{\mathbf{R}}$的对角线元素，得到矩阵$\widehat{\mathbf{R}}^*$，解得因子载荷矩阵的估计$\widehat{\mathbf{A}} = (\widehat{a}_{ij})$和特殊因子方差的估计
$$\widehat{\psi}_i = 1 - \sum_{j=1}^{k} \widehat{a}_{ij}^2, \quad i = 1, \cdots, p.$$

- 如果$\widehat{\psi}_i$全为非负，则主因子解$\widehat{\mathbf{A}}=(\widehat{a}_{ij})$是允许的，否则需考虑$h_i$的其他替代，其中常被用来替代$h_i$的简单值有以下几种：
  1. 取$\widehat{h}_i=1$，则主因子解与主成分解等价；
  2. 取$\widehat{h}_i=R_i^2$，其中R_i^2为X_i与其他所有的原变量X_j的复相关系数的平方，即X_i对其余的$p-1$个X_j的回归方程的判定系数
  3. 取$\widehat{h}_i=\max_{j\neq i}|r_{ij}|$
  4. 取$\widehat{h}_i=\dfrac{1}{p-1}\sum_{j\neq i} r_{ij}$
  5. 取$\widehat{h}_i=1/r^{ii}$，其中r^{ii}是$\mathbf{R}^{-1}$的对角元素

主因子法迭代算法如下：

步骤1： 给定特殊因子方差一个初值$\mathbf{\Psi}(0)$；

步骤2： 对$\mathbf{R}^* = \mathbf{R} - \mathbf{\Psi}(0)$ 进行主成分分解，并获得估计$\widehat{\mathbf{A}}$；

步骤3： 令$\mathbf{\Psi}(0) = \mathrm{diag}(\mathbf{R} - \widehat{\mathbf{A}}\widehat{\mathbf{A}}')$，重复步骤2，直到收敛为止。

例

取$k=2$，特殊因子方差的估计值$\widehat{\psi}_i$为：

$$0.12,\ 0.15,\ 0.09,\ 0.05,\ 0.07.$$

采用主因子法求解美国债券指数因子分析模型的载荷矩阵和公共因子方差等指标。

- 解：采用两种方法求解因子模型的因子载荷矩阵：迭代的主因子法和非迭代的主因子法。
- 编写迭代的主因子法的函数fa.ipfm()
- 编写非迭代的主因子法的函数fa.nipfm()

主因子法：迭代算法函数fa.ipfm()

```
fa.ipfm = function(R, k, d){
  p = nrow(R); diag_R = diag(R); sum_rank = sum(diag_R)
  rowname = rownames(R); colname = paste('Factor',1:k,sep='')
  A = matrix(0,nrow=p,ncol=k,dimnames=list(rowname,colname))
  mmax = 200; m = 1; h = diag_R-d
  repeat{
    diag(R) = h; h1 = h; eig = eigen(R)
    for (i in 1:k)
     A[,i] = sqrt(eig$values[i])*eig$vectors[,i]
     h = diag(A%*%t(A))
     if ((sqrt(sum((h-h1)^2))<1e-4)|m==mmax) break
     m = m+1
  }
  rowname = c('SS loadings','Proportion Var','Cumulative Var')
  B = matrix(0,nrow=3,ncol=k,dimnames=list(rowname,colname))
  for (i in 1:k){
    B[1,i] = sum(A[,i]^2); B[2,i] = B[1,i]/sum_rank
    B[3,i] = sum(B[1,1:i])/sum_rank
  }
  A = round(A,3); B = round(B,3); h = round(h,3)
  diag_R = round(diag_R, 3); method = c('Iterative Principal Factor Method')
list(method=method,loadings=A,var=cbind(common=h,spcific=diag_R-h), B = B, iterative = m)
}
```

```
d = c(0.12, 0.15, 0.09, 0.05, 0.07)
fa.pfm.fit = fa.ipfm(R, k=2, d)
#### 输出结果:
$method
[1] "Iterative Principal Factor Method"
$loadings                                  $var
     Factor1 Factor2                            common  spcific
30年  -0.958  -0.261                       30年  0.986   0.014
20年  -0.960  -0.233                       20年  0.975   0.025
10年  -0.945  -0.058                       10年  0.896   0.104
5年   -0.972   0.309                       5年   1.039  -0.039
1年   -0.770   0.297                       1年   0.681   0.319
$B
                Factor1 Factor2
SS loadings       4.268   0.310
Proportion Var    0.854   0.062
Cumulative Var    0.854   0.915
$iterative
[1] 117
```

主因子法

- 与主成分法求得的因子载荷类似，其第一个因子体现了“一般的美国债券收益率"情况
- 第二个因子的因子载荷与主成分法求得的因子载荷的符号相反，体现了第二个因子与期限呈现负相关，随期限的减少而增大，但仍体现出“期限"效应
- 主因子法用了117次迭代才得到稳定解

主因子法：非迭代算法函数fa.nipfm()

```
fa.nipfm = function(R, k){
  p = nrow(R); diag_R = diag(R); sum_rank = sum(diag_R)
  rowname = rownames(R); colname = paste('Factor',1:k,sep='')
  A = matrix(0,nrow=p,ncol=k,dimnames=list(rowname,colname))
  h = apply(abs(R-diag(1,nrow=p,ncol=p)),2,max)
  diag(R) = h; h1 = h;
  eig = eigen(R)
    for (i in 1:k){
      A[,i] = sqrt(eig$values[i])*eig$vectors[,i]
      h = diag(A%*%t(A))}
  rowname = c('SS loadings','Proportion Var','Cumulative Var')
  B = matrix(0,nrow=3,ncol=k,dimnames=list(rowname,colname))
  for (i in 1:k){
    B[1,i] = sum(A[,i]^2); B[2,i] = B[1,i]/sum_rank
    B[3,i] = sum(B[1,1:i])/sum_rank
  }
  A = round(A,3); B = round(B,3); h = round(h,3)
 diag_R = round(diag_R,3); method = c('Non-iterative Principal Factor Method')
 list(method=method, loadings=A, var=cbind(common=h, spcific=diag_R-h), B=B)
}
```

```
fb.fit = fa.nipfm(R, k=2)
#### 输出结果:
$method
[1] "Non-iterative Principal Factor Method"
$loadings
     Factor1 Factor2
30年  -0.957   0.247
20年  -0.961   0.229
10年  -0.950   0.076
5年   -0.941  -0.181
1年   -0.800  -0.448

$var
     common  spcific
30年  0.976    0.024
20年  0.975    0.025
10年  0.908    0.092
5年   0.918    0.082
1年   0.840    0.160

$B
                 Factor1 Factor2
SS loadings        4.266   0.352
Proportion Var     0.853   0.070
Cumulative Var     0.853   0.924
```

- 实际应用中，$\boldsymbol{X}$的协方差矩阵和相关系数矩阵都是未知的，都需要由样本的观测值来估计
- 因此，如果$\boldsymbol{X}$的分布已知，载荷矩阵的估计也可由极大似然估计法得到
- 假设$\boldsymbol{X}$服从p元正态分布$N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$，其中
  - ▷ $\boldsymbol{\Sigma} = \mathbf{A}\mathbf{A}' + \boldsymbol{\Psi} > 0$
  - ▷ $\mathbf{A}$为$p \times k$的公共因子载荷矩阵
  - ▷ k为预先指定的公共因子的个数，且$k < p$

- 记$\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$为来自该总体$\boldsymbol{X}$的一组简单随机样本
- 则有

$$(n-1)\mathbf{S} = \sum_{i=1}^{n} (\boldsymbol{x}_i - \bar{\boldsymbol{x}})(\boldsymbol{x}_i - \bar{\boldsymbol{x}})' \sim W_p(n-1, \boldsymbol{\Sigma}).$$

- 注意到：$\mathbf{S}$是$\mathbf{A}$和$\mathbf{\Psi}$ 的充分统计量，则可由Wishart分布的密度函数得到$\mathbf{A}$和$\mathbf{\Psi}$ 的似然函数为：

$$\begin{aligned} L(\mathbf{A}, \mathbf{\Psi}) =& c|\mathbf{S}|^{(n-p-2)/2}|\mathbf{A}\mathbf{A}' + \mathbf{\Psi}|^{-(n-1)/2} \\ & \times \exp\left\{-\frac{n-1}{2}\mathrm{tr}\left(\mathbf{S}[\mathbf{A}\mathbf{A}' + \mathbf{\Psi}]^{-1}\right)\right\} \end{aligned}$$

- 可得对数似然函数为：

$$l(\mathbf{A},\boldsymbol{\Psi})=\ln c+\frac{n-p-2}{2}\ln|\mathbf{S}|-\frac{n-1}{2}\ln|\mathbf{A}\mathbf{A}'+\boldsymbol{\Psi}| \\ -\frac{n-1}{2}\mathrm{tr}\left(\mathbf{S}[\mathbf{A}\mathbf{A}'+\boldsymbol{\Psi}]^{-1}\right).$$

- **极大似然法**就是对对数似然函数$l(\mathbf{A},\boldsymbol{\Psi})$关于$\mathbf{A}$和$\boldsymbol{\Psi}$求偏导，并令其等于零，进行求解

引理10.2.1

设矩阵$\mathbf{B}(t)$是关于标量t的一个函数矩阵，对称且可逆，则

$$\frac{\partial}{\partial t}\ln|\mathbf{B}(t)|=\mathrm{tr}\left(\mathbf{B}^{-1}(t)\frac{\partial\mathbf{B}(t)}{\partial t}\right),\qquad \frac{\partial\mathbf{B}^{-1}(t)}{\partial t}=\mathbf{B}^{-1}(t)\frac{\partial\mathbf{B}(t)}{\partial t}\mathbf{B}^{-1}(t).$$

- 对$l(\mathbf{A},\mathbf{\Psi})$分别关于$\mathbf{\Psi}$中的每个元素$\psi_i$和矩阵$\mathbf{A}$的每个元素$a_{ij}$求偏导，令其等于0，建立方程组：

$$\begin{cases} \dfrac{\partial l(\mathbf{A},\mathbf{\Psi})}{\partial \psi_i} = 0, \\ \dfrac{\partial l(\mathbf{A},\mathbf{\Psi})}{\partial a_{ij}} = 0, \quad i = 1,\cdots,p; \ \ j = 1,\cdots,k. \end{cases}$$

- 由引理10.2.1，可得

$$\begin{cases} \mathbf{\Psi} = \mathbf{S} - \mathbf{A}\mathbf{A}', \\ \mathbf{S}[\mathbf{A}\mathbf{A}' + \mathbf{\Psi}]^{-1}\mathbf{A} = \mathbf{A}. \end{cases}$$

- 方程组关于$\mathbf{A}$和$\mathbf{\Psi}$没有显式解，需要用迭代算法完成

- 下面介绍软件中常用的一种迭代算法，由下面等式

$$(\mathbf{AA}' + \mathbf{\Psi})^{-1} = \mathbf{\Psi}^{-1} - \mathbf{\Psi}^{-1}\mathbf{A}(\mathbf{I}_k + \mathbf{A}'\mathbf{\Psi}^{-1}\mathbf{A})^{-1}\mathbf{A}'\mathbf{\Psi}^{-1}$$

- 把上式代入方程组第2式，则方程组为：

$$\begin{cases} \mathbf{\Psi} = \mathbf{S} - \mathbf{AA}', \\ \mathbf{S\Psi}^{-1}\mathbf{A}(\mathbf{I}_k + \mathbf{A}'\mathbf{\Psi}^{-1}\mathbf{A})^{-1} = \mathbf{A}. \end{cases}$$

- 第2式又等价于：

$$\mathbf{S\Psi}^{-1}\mathbf{A} = \mathbf{A}(\mathbf{I}_k + \mathbf{A}'\mathbf{\Psi}^{-1}\mathbf{A}).$$

- 进一步计算，第2式为：

$$\mathbf{A}(\mathbf{A}'\mathbf{\Psi}^{-1}\mathbf{A}) = (\mathbf{S} - \mathbf{\Psi})\mathbf{\Psi}^{-1}\mathbf{A}.$$

- 用$\mathbf{\Psi}^{-1/2}$左乘上式的两边，得

$$\left[\mathbf{\Psi}^{-1/2}(\mathbf{S}-\mathbf{\Psi})\mathbf{\Psi}^{-1/2}\right]\mathbf{\Psi}^{-1/2}\mathbf{A}=\mathbf{\Psi}^{-1/2}\mathbf{A}(\mathbf{A}'\mathbf{\Psi}^{-1}\mathbf{A}).$$

- 由于因子模型的因子载荷矩阵不唯一，因此，为了使方程组有唯一的解，限定$\mathbf{A}$和$\mathbf{\Psi}$满足如下约束条件：

$$\mathbf{A}'\mathbf{\Psi}^{-1}\mathbf{A}=\mathbf{\Delta}\quad \text{为对角矩阵}.$$

- 因此，$\mathbf{\Delta}$ 的对角线元素$\delta_1,\cdots,\delta_k$为

$$\mathbf{\Psi}^{-1/2}(\mathbf{S}-\mathbf{\Psi})\mathbf{\Psi}^{-1/2}$$

的前k个特征值。

极大似然估计的迭代算法：

步骤1： 给$\boldsymbol{\Psi}$的初值$\boldsymbol{\Psi}_0$，初值的给定见主因子法；

步骤2： 把初值$\boldsymbol{\Psi}_0$代入$\boldsymbol{\Psi}^{-1/2}\mathbf{S}\boldsymbol{\Psi}^{-1/2}$，求其前$k$个特征值$\lambda_1, \cdots, \lambda_k$，以及相应的特征向量$\boldsymbol{q}_1, \cdots, \boldsymbol{q}_k$，具体为：

1. 令$\mathbf{Q} = (\boldsymbol{q}_1, \cdots, \boldsymbol{q}_k)$，$\widehat{\boldsymbol{\Delta}} = \mathrm{diag}(\lambda_1 - 1, \cdots, \lambda_k - 1)$；
2. 求解，得$\widehat{\mathbf{A}} = \boldsymbol{\Psi}_0^{1/2}\mathbf{Q}\widehat{\boldsymbol{\Delta}}^{1/2}$和$\widehat{\boldsymbol{\Psi}} = \mathrm{diag}(\mathbf{S} - \widehat{\mathbf{A}}\widehat{\mathbf{A}}')$，并替代步骤1中$\boldsymbol{\Psi}$的初值；

步骤3： 重复步骤1和步骤2，直到$\widehat{\mathbf{A}}$和$\widehat{\boldsymbol{\Psi}}$的元素收敛。

- 在R语言中，用函数factanal()进行因子载荷矩阵的极大似然估计，函数的调用格式为：

```
factanal(x, factors, data = NULL, covmat = NULL,
         n.obs = NA, subset, na.action, start = NULL,
         scores = c("none", "regression", "Bartlett"),
         rotation = "varimax", control = NULL, ...)
```

其中x为数据的公式，或者是由数据(每个样本按行输入)构成的矩阵，或者是数据框；factors是因子的个数；data是数据框，当x由公式形式输入时使用；covmat是样本的协方差矩阵或样本的相关系数矩阵，此时不必输入x；scores表示因子得分的方法(scores="regression" 表示用回归方法计算因子得分；scores="Bartlett"表示用Bartlett计算因子得分；默认为"none"，即不计算因子得分)；rotation表示旋转，默认为方差最大旋转，当rotation="none" 时，不做旋转变换。

- 例：极大似然法求解美国债券指数因子分析模型的载荷矩阵和公共因子方差等指标。

```
fa.mle.fit=factanal(factors=2,covmat=R,rotation="none",n.obs=696)
#### 输出结果:
Call:
factanal(factors = 2, covmat = R, n.obs = 696, rotation = "none")
Uniquenesses:
 30年   20年   10年    5年    1年
0.014 0.025 0.098 0.005 0.278
Loadings:
     Factor1 Factor2
30年  0.931   0.347
20年  0.935   0.316
10年  0.941   0.128
5年   0.981  -0.181
1年   0.807  -0.266
               Factor1 Factor2
SS loadings      4.239   0.340
Proportion Var   0.848   0.068
Cumulative Var   0.848   0.916
Test of the hypothesis that 2 factors are sufficient.
The chi square statistic is 6.38 on 1 degree of freedom.
The p-value is 0.0115
```

因子旋转

- 由于因子载荷矩阵是不唯一的，所以可对因子载荷矩阵进行旋转。
- 旋转的目的：使因子载荷矩阵的结构简化，使载荷矩阵每列或行的元素平方值向0和1两极分化。
- 三种常用的正交旋转法：
  1. 方差最大(varimax)法
  2. 四次方最大法
  3. 等量最大法

- 方差最大法的直观意义：希望通过因子旋转，使得每个因子上的载荷尽量拉开距离，一部分的载荷趋于1，另一部分趋于0。
- 旋转后的因子载荷矩阵为：

$$\mathbf{B} = \mathbf{A}\mathbf{G} = (b_{ij})_{p\times k},$$

其中

- ▷ $\mathbf{A}$是未旋转的$p \times k$的载荷矩阵
- ▷ $\mathbf{G}$是$k \times k$的正交矩阵，且$\mathbf{G}\mathbf{G}' = \mathbf{G}'\mathbf{G} = \mathbf{I}_k$
- ▷ b_{ij}为旋转后第i个变量X_i在第j个因子上的载荷

- 因此，方差最大准则下要求最大化的函数为：

$$\phi = \sum_{j=1}^{k} \sum_{i=1}^{p} (d_{ij}^2 - \overline{d}_j)^2,$$

这里

- ▷ $d_{ij} = \dfrac{b_{ij}}{\sqrt{h_i}}, \quad \overline{d}_j = \dfrac{1}{p} \sum_{i=1}^{p} d_{ij}^2,$
- ▷ $h_i = \sum_{j=1}^{k} b_{ij}^2$为变量$X_i$的共同度

- Kaiser (1958)给出了寻找**G**的一个迭代算法。当$k=2$时，则**G** 定义为

$$\mathbf{G}=\begin{pmatrix}\cos(\theta) & \sin(\theta)\\ -\sin(\theta) & \cos(\theta)\end{pmatrix}.$$

  - 表示坐标轴顺时针旋转θ角度，该旋转完全由θ决定
  - 将其代入最大化的函数ϕ，解出最大化的θ^*
- 当$k>2$时，就先对其中两个因子用上述方法做旋转，然后将新的旋转因子和第三个原始因子做旋转，循环迭代
  - 具体的算法参考Kaiser(1958)，与张尧庭和方开泰(2003)

例

Lawley和Maxwell (1971)提供了220名男学生的6门课程(盖尔语、英语、历史、算数、代数、几何)考试成绩的样本相关系数矩阵：

$$\mathbf{R} = \begin{pmatrix} 1.000 & & & & & \\ 0.439 & 1.000 & & & & \\ 0.410 & 0.351 & 1.000 & & & \\ 0.288 & 0.354 & 0.164 & 1.000 & & \\ 0.329 & 0.320 & 0.190 & 0.595 & 1.000 & \\ 0.248 & 0.329 & 0.181 & 0.470 & 0.464 & 1.000 \end{pmatrix}.$$

考虑极大似然法求得两因子分析模型的因子载荷，和因子载荷的旋转。

- 利用函数factanal()计算两个因子分析模型的因子载荷极大似然估计

```
names = c('盖尔语', '英语', '历史', '算数', '代数', '几何')
R.score = matrix(c(1.0, .439, .410, .288, .329, .248, .439,
                   1.0, .351, .354, .320, .329, .410, .351,
                   1.0, .164, .190, .181, .288, .354, .164,
                   1.0, .595, .470, .329, .320, .190, .595,
                   1.0, .464, .248, .329, .181, .470, .464, 1.0),
          nrow=6, ncol=6, dimnames=list(names,names), byrow=T)
fa.fit=factanal(factors=2,covmat=R.score,rotation="none",n.obs=220)
#### 输出结果:
Call:
factanal(factors=2, covmat=R.score, n.obs=220, rotation="none")
Uniquenesses:
盖尔语  英语   历史   算数   代数   几何
0.510  0.594  0.644  0.377  0.431  0.628
```

因子旋转

```
Loadings:
        Factor1 Factor2
盖尔语   0.553   0.429
英语     0.568   0.288
历史     0.392   0.450
算数     0.740  -0.273
代数     0.724  -0.211
几何     0.595  -0.132
                Factor1 Factor2
SS loadings       2.209   0.606
Proportion Var    0.368   0.101
Cumulative Var    0.368   0.469
Test of the hypothesis that 2 factors are sufficient.
The chi square statistic is 2.33 on 4 degrees of freedom.
The p-value is 0.674
```

- 所有课程在第一个因子上的载荷都是正的
- 课程盖尔语、英语和历史在第二个因子上的载荷是正的，课程算数、代数和几何在第二个因子上的载荷是负的
- Lawley和Maxwell(1971)把第一个因子称为“智商"因子
- 第二个因子不容易识别，不过他们发现语言类课程成绩高的学生第二个因子的因子得分就高，而数学类课程成绩高的学生第二个因子的因子得分低，于是称第二个因子为“数学–非数学"因子

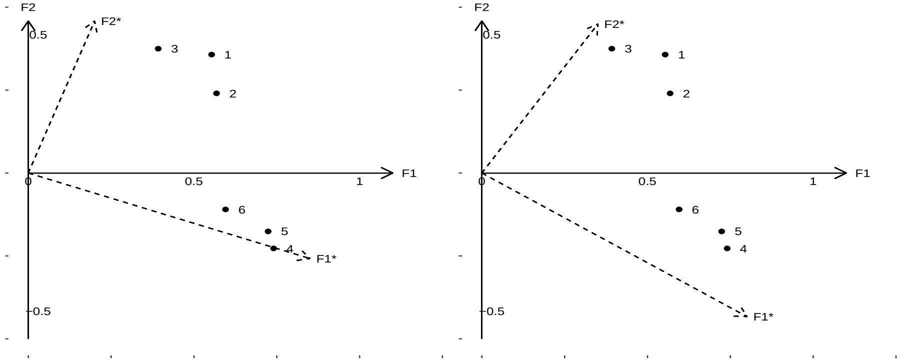


Figure: 6门课程成绩的因子旋转，左图：顺时针旋转$\theta = 20°$；右图：方差最大旋转，即顺时针旋转$\theta = 32.6°$。

- 对坐标轴顺时针旋转$\theta = 20^\circ$后，算数、代数和几何在第一个因子上的载荷较大，而在第二因子上的载荷非常小
- 因此，可认为旋转后的第一个因子为数学能力因子
- 盖尔语、英语和历史在第二个因子上的载荷相对于第一个因子上的相应载荷都稍大些，因此第二个因子解释为语言能力因子略显牵强

因子旋转

- 旋转20°后的载荷矩阵计算程序

```
A = fa.fit$loadings[1:6,]
phi = 20/180*pi
G = matrix(c(cos(phi),-sin(phi),sin(phi),cos(phi)),
           nrow=2, ncol=2)
> round(A%*%G, 3)
          [,1]   [,2]
盖尔语 0.373  0.592
英语    0.435  0.465
历史    0.215  0.557
算数    0.789 -0.003
代数    0.752  0.049
几何    0.604  0.080
```

- 载荷矩阵的方差最大旋转可应用函数varimax()实现，并计算最大旋转角度

```
T.fit = varimax(A, normalize=TRUE)
theta = acos(T.fit$rotmat[1,1])*180/pi     ## 旋转角度
> T.fit
$loadings
Loadings:
       Factor1 Factor2
盖尔语 0.235   0.659
英语   0.323   0.549
历史           0.590
算数   0.771   0.170
代数   0.724   0.213
几何   0.572   0.210

               Factor1 Factor2
SS loadings      1.612   1.203
Proportion Var   0.269   0.201
Cumulative Var   0.269   0.469
$rotmat
           [,1]      [,2]
[1,]  0.8420397 0.5394156
[2,] -0.5394156 0.8420397
```

```
> theta
[1] 32.64386
```

因子旋转

Table: 极大似然法下旋转前和旋转后因子载荷对比结果

变量	因子载荷		旋转后的因子载荷		公共方差	特殊方差
	F_1	F_2	F_1^*	F_2^*		
盖尔语	0.553	0.429	0.235	**0.659**	0.490	0.510
英语	0.568	0.288	0.323	**0.549**	0.406	0.594
历史	0.392	0.450	0.000	**0.590**	0.356	0.644
算数	0.740	−0.273	**0.771**	0.170	0.623	0.377
代数	0.724	−0.211	**0.724**	0.213	0.568	0.431
几何	0.595	−0.132	**0.572**	0.210	0.372	0.628
贡献	2.209	0.606	1.612	1.203	旋转矩阵	
贡献率	0.368	0.101	0.269	0.201	0.842	0.539
累计贡献率	0.368	0.469	0.269	0.469	−0.539	0.842

- 对美国债券指数因子模型的因子载荷进行方差最大旋转，并进行比较
- 对主成分法、主因子法和极大似然法得到的输出结果，利用函数varimax()进行计算

```
vm.pcm = varimax(fa.pcm.fit$loadings, normalize=F)
vm.ipfm = varimax(fa.pfm.fit$loadings, normalize=F)
vm.mlm = factanal(factors=2, covmat=R,
                  rotation="varimax", n.obs=696)
```

Table: 主成分法下旋转前和旋转后因子载荷对比结果

变量	因子载荷 F_1	因子载荷 F_2	旋转后的因子载荷 F_1^*	旋转后的因子载荷 F_2^*	公共方差	特殊方差
30年	−0.952	0.258	−0.921	−0.353	0.973	0.027
20年	−0.956	0.242	−0.915	−0.368	0.973	0.027
10年	−0.960	0.129	−0.851	−0.461	0.937	0.063
5年	−0.956	−0.156	−0.681	−0.688	0.938	0.062
1年	−0.825	−0.547	−0.345	−0.928	0.980	0.020
贡献	4.337	0.465	2.944	1.807	旋转矩阵	
贡献率	0.867	0.09	0.599	0.361	0.808	0.589
累计贡献率	0.867	0.960	0.599	0.960	−0.589	0.808

Table: 迭代主因子法下旋转前和旋转后因子载荷对比结果

变量	因子载荷		旋转后的因子载荷		公共方差	特殊方差
	F_1	F_2	F_1^*	F_2^*		
30年	−0.958	−0.261	−0.903	0.414	0.986	0.014
20年	−0.960	−0.233	−0.886	0.437	0.975	0.025
10年	−0.945	−0.058	−0.762	0.561	0.896	0.104
5年	−0.972	0.309	−0.548	0.860	1.039	−0.039
1年	−0.770	0.297	−0.401	0.722	0.681	0.319
贡献	4.268	0.310	2.642	1.938	旋转矩阵	
贡献率	0.854	0.062	0.528	0.388	0.767	−0.641
累计贡献率	0.854	0.915	0.528	0.916	0.641	0.767

Table: 极大似然法下旋转前和旋转后因子载荷对比结果

变量	因子载荷		旋转后的因子载荷		公共方差	特殊方差
	F_1	F_2	F_1^*	F_2^*		
30年	0.931	0.347	0.895	0.431	0.987	0.014
20年	0.935	0.316	0.876	0.455	0.975	0.025
10年	0.941	0.128	0.744	0.590	0.902	0.098
5年	0.981	−0.181	0.549	0.833	0.995	0.005
1年	0.807	−0.266	0.367	0.766	0.722	0.278
贡献	4.239	0.340	2.557	2.022	旋转矩阵	
贡献率	0.848	0.068	0.511	0.404	0.716	0.698
累计贡献率	0.848	0.916	0.511	0.916	0.698	−0.716

因子旋转

1. 旋转后的第一个因子的载荷绝对值随期限递增，而第二个因子的载荷绝对值随期限递减
2. 旋转前所有变量在第一个因子的载荷都很大也都很接近，所以第一因子体现了债券收益率的总效应，第二个因子随时间递减，体现了时间效应
3. 对比结果，旋转前因子实际意义更清晰

- 问题：公共因子的个数k是否能充分刻画变量的协方差矩阵或相关系数矩阵？
- 考虑如下的检验问题：

$$H_0: \ \mathbf{\Sigma} = \mathbf{A}\mathbf{A}' + \mathbf{\Psi},$$

这里，假设$\mathbf{A}$为$p \times k$的矩阵。

- 假设$X=(X_1,\cdots,X_p)'$服从p元正态分布，应用似然比检验方法，可构造Bartlett修正的检验统计量为：

$$\chi^2=\left(n-1-\frac{1}{6}(2p+5)-\frac{2}{3}k\right)\ln\frac{|\widehat{\mathbf{A}}\widehat{\mathbf{A}}'+\widehat{\mathbf{\Psi}}|}{|\mathbf{S}|},$$

其中

  - $\widehat{\mathbf{A}}$和$\widehat{\mathbf{\Psi}}$为极大似然估计
  - $\mathbf{S}=\dfrac{1}{n-1}\sum_{i=1}^{n}(\boldsymbol{x}_i-\overline{\boldsymbol{x}})(\boldsymbol{x}_i-\overline{\boldsymbol{x}})'$ 为样本协方差矩阵

- 当$n\to\infty$时，统计量χ^2的渐近分布为：

$$\chi^2\xrightarrow{d}\chi^2_f,$$

其中自由度$f=\left((p-k)^2-p-k\right)/2$。

- 因此，当$\chi^2 > \chi_f^2(\alpha)$时，拒绝原假设$H_0$，认为需要引入更多因子，否则接受原假设$H_0$。
- 特别地，当$k=0$时，原假设H_0就是假设$\boldsymbol{\Sigma}$为对角矩阵，该检验问题就等价于$\boldsymbol{X}$的独立性检验。
- Lawley (1940) 进一步提出下面的检验统计量：

$$\chi^2_{\text{Lawley}} = \left(n-1-\frac{1}{6}(2p+5)-\frac{2}{3}k\right)\sum\sum_{i<j}\frac{(s_{ij}-\widehat{\sigma}_{ij})^2}{\widehat{\psi}_i\widehat{\psi}_j}.$$

  - ▷ s_{ij}为样本协方差矩阵$\mathbf{S}$的第(i,j)个元素
  - ▷ $\widehat{\sigma}_{ij}=\sum_{h=1}^{k}\widehat{a}_{ih}\widehat{a}_{jh},\quad i\neq j,$
  - ▷ $\widehat{a}_{ih}$为$\widehat{\mathbf{A}}$的第(i,h)个元素

因子分析模型的拟合优度检验

- 可以证明：$\chi^2_{\text{Lawley}} \xrightarrow{d} \chi^2_f$
- 在实际应用中，需要从小到大逐一来指定k，用以上方法检验，直到检验H_0显著为止来确定k的大小
- 由于χ^2统计量的自由度f一定是正的，故要求指定的公共因子的个数满足如下条件：

$$k < \frac{1}{2}\left(2p + 1 - \sqrt{8p + 1}\right).$$

- 关于因子个数的检验，可见函数factanal()的输出结果

- **AIC或BIC信息量准则：** 公共因子个数k也可通过极小化AIC或BIC信息量准则进行选取
- **AIC信息量准则：**

$$\text{AIC}(k) = -2\ln L(k) + 2[p(k+1) - k(k-1)/2],$$

其中

- ▷ $L(k)$表示包含k个公共因子的因子模型下的极大似然函数
- ▷ 在正态分布假设下，定义为：$-2\ln L(k) = np\ln(2\pi) + n\ln|\widehat{\Sigma}_k| + \text{tr}\left(\widehat{\Sigma}_k^{-1}\mathbf{S}\right)$
- ▷ $\widehat{\Sigma}_k$表示包含k个公共因子的因子模型的协方差矩阵的估计

- k的选择为：$\widehat{k} = \arg\min\limits_{k} \text{AIC}(k)$

- **BIC信息量准则：**

$$\mathsf{BIC}(k) = -2\ln L(k) + \ln(n)[p(k+1) - k(k-1)/2],$$

其中

- ▷ $L(k)$表示包含k个公共因子的因子模型下的极大似然函数
- ▷ 在正态分布假设下，定义为：$-2\ln L(k) = np\ln(2\pi) + n\ln|\widehat{\Sigma}_k| + \mathrm{tr}\left(\widehat{\Sigma}_k^{-1}\mathbf{S}\right)$
- ▷ $\widehat{\Sigma}_k$表示包含k个公共因子的因子模型的协方差矩阵的估计

- k的选择为：$\widehat{k} = \arg\min_k \mathsf{BIC}(k)$

例

对美国债券指数因子模型做检验，即考虑检验问题：

$$H_0 : \mathbf{\Sigma} = \mathbf{A}\mathbf{A}' + \mathbf{\Psi},$$

其中取$k = 2$和显著性水平$\alpha = 0.05$。

解：由前面计算可知：

$$\widehat{\mathbf{A}}\widehat{\mathbf{A}}' + \widehat{\mathbf{\Psi}} = \begin{pmatrix} 1.00 & & & & \\ 0.98 & 1.0 & & & \\ 0.92 & 0.92 & 1.00 & & \\ 0.85 & 0.86 & 0.90 & 1.00 & \\ 0.66 & 0.67 & 0.73 & 0.84 & 1.00 \end{pmatrix}.$$

- 于是，有
$$\frac{|\widehat{\mathbf{A}}\widehat{\mathbf{A}}' + \widehat{\mathbf{\Psi}}|}{|\mathbf{S}|} = \frac{0.00032}{0.00029} = 1.087.$$

- 计算自由度为$f = ((5-2)^2 - 5 - 2)/2 = 1$，且$\chi^2$统计量为：
$$\left(696 - 1 - \frac{2\times 5 + 5}{6} - \frac{2\times 2}{3}\right)\ln(1.087) = 57.658$$
$$> \chi_1^2(0.05) = 3.84.$$

- 因此，在显著性水平为$\alpha = 0.05$下，认为两因子分析模型来描述美国债券指数数据不够充分。

- 在R语言中，可使用程序包psych中的函数fa.parallel()绘制碎石图，直观确定因子个数k

```
library(psych); fa.parallel(R, n.obs=696)
```

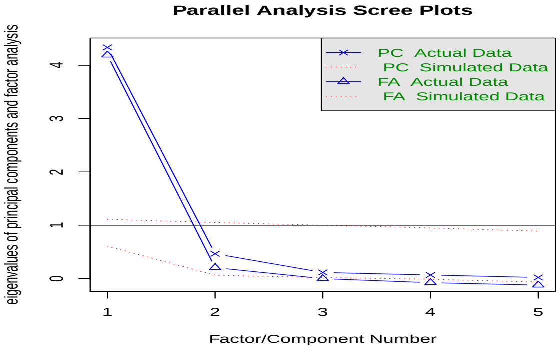


Figure: 平行分析的碎石图，其中“PC"表示主成分分析，“FA"表示因子分析。

因子得分

- 应用中，需要估计公共因子$\boldsymbol{F}$，给出公共因子的值，即因子得分。
- 因子得分，主要有两种：
  1. Thomson因子得分
  2. Bartlett因子得分

- 因子分析模型中$\boldsymbol{X}$和$\boldsymbol{F}$都是随机的，Thomson (1935) 提出的基于Bayes方法的因子得分。
- 假设
  - ▷ 因子分析模型中的向量$\boldsymbol{X}$服从p元正态分布$N_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$
  - ▷ 公共因子$\boldsymbol{F}$的先验分布为$N_k(\mathbf{0}, \mathbf{I}_k)$
- 由多元正态分布的性质，可知：

1. $\boldsymbol{X}-\boldsymbol{\mu}$和$\boldsymbol{F}$的联合分布为正态分布，即

$$\begin{pmatrix} \boldsymbol{X}-\boldsymbol{\mu} \\ \boldsymbol{F} \end{pmatrix} \sim N_{p+k}(\mathbf{0}, \boldsymbol{\Sigma}^*), \quad \text{其中} \quad \boldsymbol{\Sigma}^* = \begin{pmatrix} \mathbf{A}\mathbf{A}' + \boldsymbol{\Psi} & \mathbf{A} \\ \mathbf{A}' & \mathbf{I}_k \end{pmatrix};$$

2. 给定$\boldsymbol{X}$，$\boldsymbol{F}$的条件分布仍为正态分布，其均值向量和协方差矩阵分别为：

$$E(\boldsymbol{F}|\boldsymbol{X}) = \mathbf{A}'(\mathbf{A}\mathbf{A}' + \boldsymbol{\Psi})^{-1}(\boldsymbol{X}-\boldsymbol{\mu}) =: \boldsymbol{F}^*,$$

$$\mathrm{Cov}(\boldsymbol{F}|\boldsymbol{X}) = \mathbf{I}_k - \mathbf{A}'(\mathbf{A}\mathbf{A}' + \boldsymbol{\Psi})^{-1}\mathbf{A}.$$

- Thomson因子得分：由于$\boldsymbol{F}$的后验分布就是给定$\boldsymbol{X}$下$\boldsymbol{F}$的条件分布，故$\boldsymbol{F}$的Bayes估计就是$\boldsymbol{F}$的后验均值向量$\boldsymbol{F}^*$，称之为Thomson因子得分。
- 下面给出Thomson因子得分的两种等价形式：

$$\boldsymbol{F}^* = \mathbf{A}'(\mathbf{A}\mathbf{A}' + \boldsymbol{\Psi})^{-1}(\boldsymbol{X} - \boldsymbol{\mu}),$$

$$\boldsymbol{F}^* = (\mathbf{I}_k + \mathbf{A}'\boldsymbol{\Psi}^{-1}\mathbf{A})^{-1}\mathbf{A}'\boldsymbol{\Psi}^{-1}(\boldsymbol{X} - \boldsymbol{\mu}).$$

- 可见：Thomson因子得分$\boldsymbol{F}^*$为给定$\boldsymbol{X}$下$\boldsymbol{F}$的条件期望，故 $(\mathbf{I}_k+\mathbf{A}'\mathbf{\Psi}^{-1}\mathbf{A})^{-1}\mathbf{A}'\mathbf{\Psi}^{-1}$ 可看作是向量$\boldsymbol{X}$对因子$\boldsymbol{F}$的回归系数
- 通过估计回归系数得到因子得分，可看作是对条件期望$E(\boldsymbol{F}|\boldsymbol{X})$ 的估计，因此也把Thomson因子得分的方法称为回归方法
- 给定观测样本$\boldsymbol{x}_1,\cdots,\boldsymbol{x}_n$，可得第$i$个样本的Thomson 因子得分为：

$$\widehat{\boldsymbol{F}}_i^* = (\mathbf{I}_k+\widehat{\mathbf{A}}'\widehat{\mathbf{\Psi}}^{-1}\widehat{\mathbf{A}})^{-1}\widehat{\mathbf{A}}'\widehat{\mathbf{\Psi}}^{-1}(\boldsymbol{x}_i-\bar{\boldsymbol{x}}), \quad i=1,\cdots,n.$$

  - ▷ $\bar{\boldsymbol{x}}$、$\widehat{\mathbf{A}}$和$\widehat{\mathbf{\Psi}}$为极大似然估计

- 同样，当$\mathbf{A}$和$\mathbf{\Psi}$的估计是基于样本相关阵$\widehat{\mathbf{R}}$，则第i个样本的Thomson因子得分为：

$$\widehat{F}_i^* = (\mathbf{I}_k + \widehat{\mathbf{A}}'\widehat{\mathbf{\Psi}}^{-1}\widehat{\mathbf{A}})^{-1}\widehat{\mathbf{A}}'\widehat{\mathbf{\Psi}}^{-1} z_i, \quad i = 1, \cdots, n,$$

其中z_i为对第i个观测样本$\boldsymbol{x}_i$进行标准化后的数据。

- 为了避免计算误差，有时也直接对$\mathbf{A}\mathbf{A}' + \mathbf{\Psi}$用样本协方差阵$\widehat{\mathbf{\Sigma}} = \mathbf{S}$来估计，得到第$i$个样本的Thomson因子得分为：

$$\widehat{F}_i^* = \widehat{\mathbf{A}}'\widehat{\mathbf{\Sigma}}^{-1}(\boldsymbol{x}_i - \bar{\boldsymbol{x}}) = \widehat{\mathbf{A}}'\mathbf{S}^{-1}(\boldsymbol{x}_i - \bar{\boldsymbol{x}}), \quad i = 1, \cdots, n.$$

- 注意到：Thomson因子得分是有偏的，即

$$E(\boldsymbol{F}^*|\boldsymbol{F}) = (\mathbf{I}_k + \mathbf{A}'\boldsymbol{\Psi}^{-1}\mathbf{A})^{-1}\mathbf{A}'\boldsymbol{\Psi}^{-1}\mathbf{A}\boldsymbol{F} \neq \boldsymbol{F}.$$

- Bartlett (1937, 1938b) 基于给定$\boldsymbol{F}$，考虑$\boldsymbol{X}$的条件分布，提出了另一种因子得分方法。
- 假设均值向量$\boldsymbol{\mu}$、载荷矩阵$\mathbf{A}$和特殊因子方差阵$\boldsymbol{\Psi}$ 已知，则因子分析模型可以写成：

$$\boldsymbol{X} - \boldsymbol{\mu} = \mathbf{A}\boldsymbol{F} + \varepsilon.$$

- 可见，因子分析模型就是一个线性回归模型，公共因子$\boldsymbol{F}$ 就可被视为待估参数。

- 假设$\boldsymbol{X}$来自p元正态分布，即$\boldsymbol{X}-\boldsymbol{\mu}\sim N_p(\mathbf{A}\boldsymbol{F},\boldsymbol{\Psi})$。
- 于是，$\boldsymbol{X}$的对数似然函数为：

$$l(\boldsymbol{F})=\frac{1}{2}[(\boldsymbol{X}-\boldsymbol{\mu})-\mathbf{A}\boldsymbol{F}]'\boldsymbol{\Psi}^{-1}[(\boldsymbol{X}-\boldsymbol{\mu})-\mathbf{A}\boldsymbol{F}]-\frac{1}{2}\ln|2\pi\boldsymbol{\Psi}|.$$

- 可得$\boldsymbol{F}$的Bartlett因子得分为：

$$\widehat{\boldsymbol{F}}=(\mathbf{A}'\boldsymbol{\Psi}^{-1}\mathbf{A})^{-1}\mathbf{A}'\boldsymbol{\Psi}^{-1}(\boldsymbol{X}-\boldsymbol{\mu}).$$

- 思考：当$\boldsymbol{X}$不服从p元正态分布时怎么办？

性质10.5.1

Bartlett因子得分具有无偏性，即$E(\widehat{\boldsymbol{F}}|\boldsymbol{F}) = \boldsymbol{F}$。

- 给定观测样本$\boldsymbol{x}_1, \cdots, \boldsymbol{x}_n$，可得第$i$个样本的Bartlett 因子得分为：

$$\widehat{\boldsymbol{F}}_i = (\widehat{\mathbf{A}}'\widehat{\boldsymbol{\Psi}}^{-1}\widehat{\mathbf{A}})^{-1}\widehat{\mathbf{A}}'\widehat{\boldsymbol{\Psi}}^{-1}(\boldsymbol{x}_i - \bar{\boldsymbol{x}}), \quad i = 1, \cdots, n.$$

- 如果$\widehat{\mathbf{A}}$和$\widehat{\boldsymbol{\Psi}}$为$\mathbf{A}$ 和$\boldsymbol{\Psi}$的极大似然估计，它们需要满足唯一性条件，即

$$\widehat{\mathbf{A}}'\widehat{\boldsymbol{\Psi}}^{-1}\widehat{\mathbf{A}} = \widehat{\boldsymbol{\Delta}},$$

其中$\widehat{\boldsymbol{\Delta}}$为对角矩阵。

- 若$\mathbf{A}$和$\mathbf{\Psi}$的估计是基于样本相关系数矩阵$\widehat{\mathbf{R}}$，则第i个样本的Bartlett因子得分为：

$$\widehat{F}_i = (\widehat{\mathbf{A}}_z'\widehat{\mathbf{\Psi}}_z^{-1}\widehat{\mathbf{A}}_z)^{-1}\widehat{\mathbf{A}}_z'\widehat{\mathbf{\Psi}}_z^{-1}z_i, \quad i = 1, \cdots, n,$$

其中

  - ▷ z_i为对第i 个观测样本$\boldsymbol{x}_i$ 进行标准化后的数据
  - ▷ $\widehat{\mathbf{R}} = \widehat{\mathbf{A}}_z\widehat{\mathbf{A}}_z' + \widehat{\mathbf{\Psi}}_z$

- 当因子载荷和特殊因子方差的估计是采用主成分法时，则第i个样本的Bartlett因子得分为：

$$\widehat{F}_i = (\widetilde{\mathbf{A}}'\widetilde{\mathbf{A}})^{-1}\widetilde{\mathbf{A}}'(\boldsymbol{x}_i - \bar{\boldsymbol{x}}), \quad i = 1, \cdots, n.$$

性质10.5.2

当因子载荷采用主成分法来估计时，则相应Bartlett因子得分$\widehat{F}$就是X的前k个样本主成分的标准化。

解：主成分法的因子载荷矩阵估计为：

$$\widetilde{\mathbf{A}}=\left(\sqrt{\widehat{\lambda}_1}\boldsymbol{u}_1,\cdots,\sqrt{\widehat{\lambda}_k}\boldsymbol{u}_k\right),$$

其中$\widehat{\lambda}_i$和$\boldsymbol{u}_i$分别为样本协方差矩阵$\mathbf{S}$的第i个特征值和相应的特征向量。

可得

$$\widehat{\boldsymbol{F}}_i = \begin{pmatrix} \frac{1}{\sqrt{\widehat{\lambda}_1}}\boldsymbol{u}_1'(\boldsymbol{x}_i - \overline{\boldsymbol{x}}) \\ \vdots \\ \frac{1}{\sqrt{\widehat{\lambda}_k}}\boldsymbol{u}_k'(\boldsymbol{x}_i - \overline{\boldsymbol{x}}) \end{pmatrix}, \qquad i = 1, \cdots, n.$$

简单计算，有：$\frac{1}{n}\sum_{i=1}^{n}\widehat{\boldsymbol{F}}_i = \mathbf{0}, \quad \frac{1}{n-1}\sum_{i=1}^{n}\widehat{\boldsymbol{F}}_i\widehat{\boldsymbol{F}}_i' = \mathbf{I}_k$。

令$z_{ij} = \boldsymbol{u}_j'(\boldsymbol{x}_i - \overline{\boldsymbol{x}})$，其中$i = 1, \cdots, n; j = 1, \cdots, k$，则$z_{ij}$表示第$i$个样本在第$j$个样本主成分的得分。$\widehat{\lambda}_j$表示第$j$个主成分的方差估计。

例

求美国债券因子分析模型中的因子得分问题。

解： 由极大似然法求得的因子载荷$\mathbf{A}_z$和特殊因子方差$\mathbf{\Psi}_z$ 分别为：

$$\widehat{\mathbf{A}}_z = \begin{pmatrix} 0.931 & 0.347 \\ 0.935 & 0.316 \\ 0.941 & 0.128 \\ 0.981 & -0.181 \\ 0.807 & -0.266 \end{pmatrix}, \quad \widehat{\mathbf{\Psi}}_z = \begin{pmatrix} 0.014 & & & & \\ & 0.025 & & & \\ & & 0.098 & & \\ & & & 0.005 & \\ & & & & 0.278 \end{pmatrix}.$$

- 对于标准化的观测值$\boldsymbol{z}=(1.4,-0.2,0.5,-1.0,0.8)'$，算得Bartlett因子得分为

$$\widehat{\boldsymbol{F}}=(\widehat{\mathbf{A}}_z'\widehat{\boldsymbol{\Psi}}_z^{-1}\widehat{\mathbf{A}}_z)^{-1}\widehat{\mathbf{A}}_z'\widehat{\boldsymbol{\Psi}}_z^{-1}\boldsymbol{z}=\begin{pmatrix}-0.338\\3.504\end{pmatrix}.$$

- Thomson因子得分为

$$\widehat{\boldsymbol{F}}^*=(\mathbf{I}_2+\widehat{\mathbf{A}}_z'\widehat{\boldsymbol{\Psi}}_z^{-1}\widehat{\mathbf{A}}_z)^{-1}\widehat{\mathbf{A}}_z'\widehat{\boldsymbol{\Psi}}_z^{-1}\boldsymbol{z}=\begin{pmatrix}-0.337\\3.335\end{pmatrix}.$$

- Bartlett因子得分是无偏的
- Thomson因子得分是有偏的
- 在预测均方误差准则下，Thomson 因子得分优于Bartlett 因子得分

性质10.5.3

在$\boldsymbol{\mu}$、$\mathbf{A}$ 和$\boldsymbol{\Psi}$已知情形下，Thomson因子得分的预测均方误差小于Bartlett因子得分的预测均方误差，即

$$\mathrm{PMSE}(\boldsymbol{F}^*) < \mathrm{PMSE}(\widehat{\boldsymbol{F}}).$$

例

考虑31个国家/地区不同年龄段(0岁、25岁、50岁和75岁)男性与女性的预期寿命数据的因子分析，该数据包含8个变量：m0、m25、m50、m75、w0、w25、w50和w75，共有248个观测值，其中“m"和“w"分别代表男性和女性。考虑极大似然法求得因子分析模型的因子载荷，和因子载荷的方差最大旋转，并计算Thomson因子得分和Bartlett因子得分，对计算结果进行合理的分析和解释。

- 首先，考虑数据的相关系数矩阵和公共因子个数k

```
library(ggplot2); library(GGally); library(psych)
life.data = read.table("life.txt")
life = structure(.Data=life.data[,2:9], class = "data.frame",
 names = c("m0", "m25", "m50", "m75", "w0", "w25", "w50", "w75"),
    row.names = life.data[,1])
ggpairs(data = life)
x = scale(life, center = T, scale = T)
k = fa.parallel(x, fa = "fa")
Parallel analysis suggests that the number of factors =  2
```

案例与R语言计算

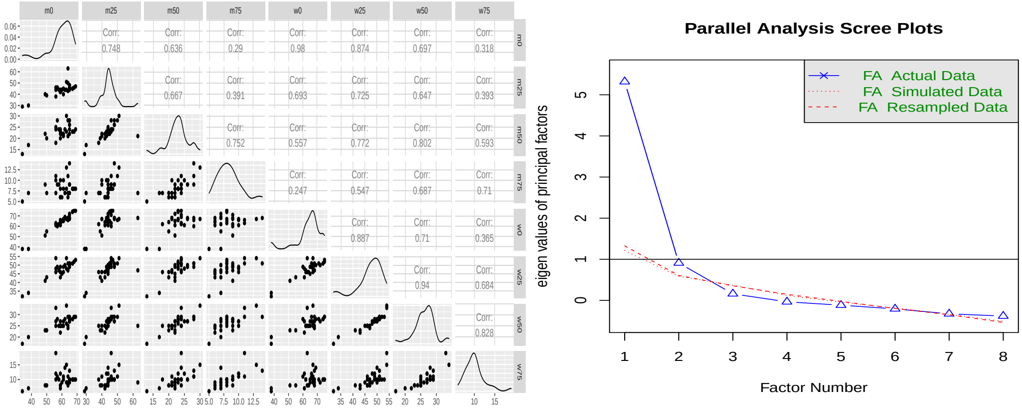


Figure: 31个国家/地区预期寿命数据的矩阵散点图和碎石图。左图：矩阵散点图；右图：平行分析的碎石图。

Table: 预期寿命数据的因子数检验

因子个数k	χ^2统计量值	自由度	p值
2	45.24	13	1.91×10^{-5}
3	6.73	7	0.458

- 综合平行分析的碎石图和拟合优度检验结果，下面考虑三因子分析模型。

```
factanal(life, factors=3, rotation="varimax")
#### 输出结果:
Call:
factanal(x = life, factors = 3, rotation = "varimax")
Uniquenesses:
   m0   m25   m50   m75    w0   w25   w50   w75
0.005 0.362 0.066 0.288 0.005 0.011 0.020 0.146
Loadings:
    Factor1 Factor2 Factor3
m0  0.964   0.122   0.226
m25 0.646   0.169   0.438
m50 0.430   0.354   0.790
m75         0.525   0.656
w0  0.970   0.217
w25 0.764   0.556   0.310
w50 0.536   0.729   0.401
w75 0.156   0.867   0.280
               Factor1 Factor2 Factor3
SS loadings      3.375   2.082   1.640
Proportion Var   0.422   0.260   0.205
Cumulative Var   0.422   0.682   0.887
```

- 第一公共因子在变量m0、w0、m25 和w25上的载荷很大，体现了婴幼儿以及青年因子(不分男女)
- 第二公共因子在变量w50和w75上的载荷较大，体现了老年女性因子
- 第三公共因子则在变量m50和m75的载荷较大，体现的是老年男性因子

```
life.Thomson=factanal(life, factors=3, scores="regression")$scores
#### 输出Thomson因子得分:
                   Factor1    Factor2     Factor3
Algeria         -0.258062561  1.90095771  1.91581631
Cameroon        -2.782495791 -0.72340014 -1.84772224
Madagascar      -2.806428187 -0.81158820 -0.01210318
Mauritius        0.141004934 -0.29028454 -0.85862443
Reunion         -0.196352142  0.47429917 -1.55046466
Seychelles       0.367371307  0.82902375 -0.55214085
South Africa(C) -1.028567629 -0.08065792 -0.65421971
South Africa(W)  0.946193522  0.06400408 -0.91995289
Tunisia         -0.862493550  3.59177195 -0.36442148
Canada           1.245304248  0.29564122 -0.27342781
Costa Rica       0.508736247 -0.50500435  1.01328707
Dominican Rep    0.106044085  0.01111171  1.83871599
```

El Salvador	-0.608155779	0.65100820	0.48836431
Greenland	0.235114220	-0.69123901	-0.38558654
Grenada	0.132008172	0.25241049	-0.15220645
Guatemala	-1.450336359	-0.67765804	0.65911906
Honduras	0.043253249	-1.85175707	0.30633182
Jamaica	0.462124701	-0.51918493	0.08032855
Mexico	-0.052332675	-0.72020002	0.44417800
Nicaragua	0.268974443	0.08407227	1.70568388
Panama	0.442333434	-0.73778272	1.25218728
Trinidad(62)	0.711367053	-0.95989475	-0.21545329
Trinidad (67)	0.787286051	-1.10729029	-0.51958264
United States (66)	1.128331259	0.16389896	-0.68177046
United States (NW66)	0.400058903	-0.36230253	-0.74299137
United States (W66)	1.214345385	0.40877239	-0.69225320
United States (67)	1.128331259	0.16389896	-0.68177046
Argentina	0.731344988	0.24811968	-0.12817725
Chile	0.009751528	0.75222637	-0.49198911
Columbia	-0.240602517	-0.29543613	0.42919600
Ecuador	-0.723451797	0.44246371	1.59164974

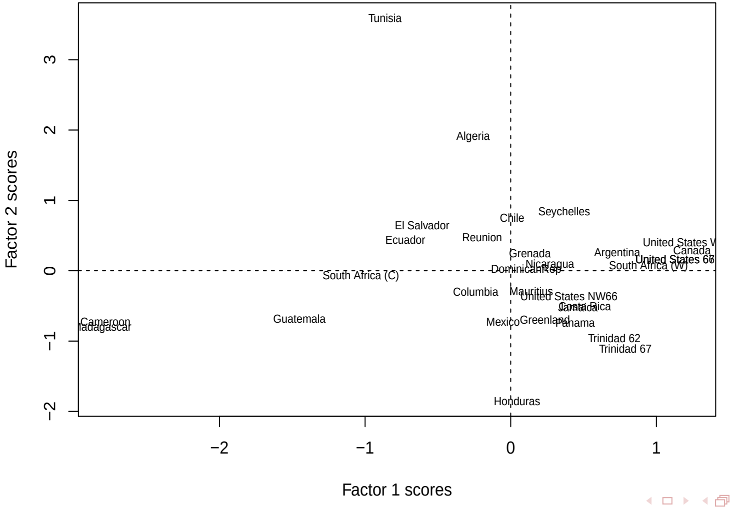
Bartlett 因子得分

```
life.Bartlett=factanal(life, factors=3, scores="Bartlett")$scores
#### 输出Bartlett因子得分:
                      Factor1     Factor2     Factor3
Algeria           -0.28185559  1.90224049  2.01985697
Cameroon          -2.77860099 -0.65968221 -1.97669190
Madagascar        -2.81776560 -0.84156031  0.04976239
Mauritius          0.14990311 -0.26324684 -0.93282800
Reunion           -0.18592771  0.57388071 -1.72931338
Seychelles         0.37106713  0.89556436 -0.65165961
South Africa(C)   -1.02799918 -0.04929805 -0.70806792
South Africa(W)    0.95847911  0.10878620 -1.02440011
Tunisia           -0.87589806  3.78955241 -0.56914746
Canada             1.25289270  0.31934241 -0.32603963
Costa Rica         0.50450230 -0.58098452  1.13647913
Dominican Rep      0.09110764 -0.07827056  2.02393717
```

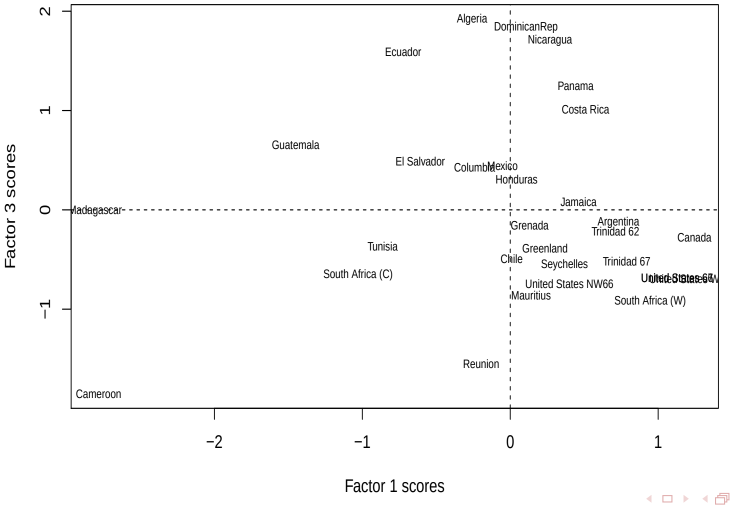
```
El Salvador          -0.61752134  0.66137065  0.51133062
Greenland             0.24186849 -0.70733470 -0.39302839
Grenada               0.13310174  0.27182730 -0.18106155
Guatemala            -1.46091026 -0.73829532  0.77121434
Honduras              0.04714225 -1.95814114  0.42727331
Jamaica               0.46553459 -0.55025834  0.10989573
Mexico               -0.05389786 -0.77717677  0.52479135
Nicaragua             0.25573169  0.00421932  1.87247870
Panama                0.43654337 -0.83665601  1.41152785
Trinidad(62)          0.72000424 -0.99913199 -0.19653474
Trinidad (67)         0.79935647 -1.13923632 -0.52499395
United States (66)    1.13920071  0.20139021 -0.76843446
United States (NW66)  0.40953762 -0.34532251 -0.80412295
United States (W66)   1.22491184  0.45855977 -0.79263233
United States (67)    1.13920071  0.20139021 -0.76843446
Argentina             0.73527845  0.26413406 -0.15941503
Chile                 0.01139502  0.81325508 -0.57865860
Columbia             -0.24442466 -0.33010210  0.48917656
Ecuador              -0.74195795  0.38918451  1.73774034
```

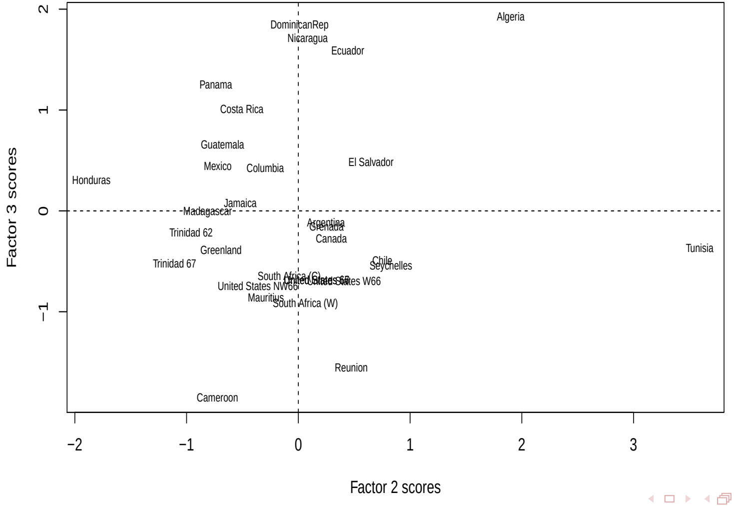
- 对比Thomson因子得分和Bartlett因子得分，两种因子得分相差不大
- 为了直观，下面用二维直角坐标来刻画两两因子的因子得分情况

```
life.df = data.frame(life, life.Thomson)
attach(life.df)
plot(Factor1,Factor2, type='n', xlab='Factor 1 scores',
     ylab='Factor 2 scores')
text(Factor1, Factor2, labels = row.names(life.df), cex = 0.7)
abline(h=0, lty=2); abline(v=0, lty=2)
```


Tunisia
Algeria
Seychelles
Chile
El Salvador
Reunion
Ecuador
Grenada
Argentina
United States V
Canada
United States 66
Nicaragua
South Africa (W)
DominicanRep
South Africa (C)
Columbia
Mauritius
United States NW66
Costa Rica
Jamaica
Cameroon
Madagascar
Guatemala
Mexico
Greenland
Panama
Trinidad 62
Trinidad 67
Honduras
3
2
1
0
−1
−2
Factor 2 scores
−2
−1
0
1
Factor 1 scores

案例与R语言计算


Algeria
DominicanRep
Nicaragua
Ecuador
Panama
Costa Rica
Guatemala
El Salvador
Columbia
Mexico
Honduras
Jamaica
Madagascar
Grenada
Argentina
Trinidad 62
Canada
Tunisia
Greenland
Chile
Seychelles
Trinidad 67
South Africa (C)
United States NW66
Mauritius
South Africa (W)
Reunion
Cameroon
2
1
0
−1
−2
−1
0
1
Factor 3 scores
Factor 1 scores


Algeria
DominicanRep
Nicaragua
Ecuador
Panama
Costa Rica
Guatemala
Mexico
Columbia
El Salvador
Honduras
Jamaica
Madagascar
Argentina
Grenada
Canada
Trinidad 62
Greenland
Tunisia
Trinidad 67
Chile
Seychelles
South Africa (C)
United States 67
United States W66
United States NW66
Mauritius
South Africa (W)
Reunion
Cameroon
2
1
0
−1
−2
3
Factor 3 scores
Factor 2 scores

- 对“因子1—因子2"的得分散点图，Madagascar(马达加斯加)和Cameroon(喀麦隆)这两个国家在散点图的左下方，说明这两国的幼儿、青年以及老年女性的平均生存时间都很低；Tunisia (突尼斯)在因子2上的得分明显高于其他国家，说明该国老年女性更长寿；
- 对“因子1—因子3"的得分散点图，Cameroon(喀麦隆)位于散点图的左下方，说明该国婴幼儿、青年以及老年男性的平均生存时间都很低，并结合“因子1—因子2" 的得分散点图，表明Cameroon(喀麦隆)的人均寿命都比别的国家要低。另外，Madagascar(马达加斯加)虽然在因子1上的得分很低，但其在因子3上处于中等水平；

- 对“因子2—因子3"的得分散点图，Tunisia (突尼斯)虽然在因子2上的得分远高于其他国家，但其在因子3上的得分并不高。Algeria (阿尔及利亚)在因子2 和因子3上得分相对比较高。说明Tunisia (突尼斯)老年女性的预期寿命远高于其他国家，而老年男性的预期寿命不算很高。Algeria (阿尔及利亚)老年人的预期寿命都很高。

- 主成分分析仅仅是一种数据变换，对数据来源的总体分布的协方差结构不做任何假设；而因子分析假设数据来源于因子分析模型，即总体分布的协方差矩阵为$\mathbf{\Sigma} = \mathbf{A}\mathbf{A}' + \mathbf{\Psi}$；
- 主成分分析强调的是从原始变量到潜在的主成分变量的变化，即主成分分析是将主成分表示为原变量的线性组合；而因子分析强调的是从潜在因子到原始变量的变换，即因子分析模型是将原始变量表示为公共因子和特殊因子的线性组合；
- 主成分与原始变量没有因果关系，而因子分析的因子与原始变量有因果关系，因子是因，表征变量是果；

因子分析与主成分分析的关系

- 主成分分析中主成分的个数和变量维数p相同，它是将一组具有相关性的原始变量变换为一组不相关的变量，而因子分析的目的是要用尽可能少的公因子，以便构造一个结构简单的因子分析模型；
- 主成分分析没有旋转，而因子分析可以旋转以使因子更好解释。因子分析的结果是不唯一的，而主成分分析的结果是唯一的；
- 主成分分析侧重“变异量"，而因子分析更重视相关变量的“共变异量"。


多元统计分析
统计与数据科学丛书 4
STATISTICS AND DATA SCIENCE SERIES 4
多元统计分析
李高荣　吴密霞　编著
Multivariate Statistical Analysis
科学出版社

谢谢，请多提宝贵意见！