多元统计

陈崇双

西南交通大学数学学院统计系

ccsmars@swjtu.edu.cn

2018-2019学年

1. 对应分析
   - 对应分析的必要性
   - 联列表
   - 基本理论

第一节：对应分析的必要性

因子分析类型：

- R型因子分析: 研究变量之间的相关关系
- Q型因子分析: 研究样本之间的相关关系

每种类型的因子分析，单独从一个角度（变量或者样本）去分析数据的信息，不能厘清之间的关系。

第一节：对应分析的必要性

因子分析类型：

- R型因子分析: 研究变量之间的相关关系
- Q型因子分析: 研究样本之间的相关关系

每种类型的因子分析，单独从一个角度（变量或者样本）去分析数据的信息，不能厘清之间的关系。对应分析结合了两种类型因子分析，同时对数据表中的行与列进行降维处理，简化了数据结构，如以二维图形表示数据表中行与列之间的关系。

对应分析的必要性

对应分析的思想，由M. W. Richardson和G. F. Kuder提出，日本统计学家Chikio Hayashi和法国统计学家Jean-Paul Benzécri发展了该方法。

M. W. Richardson, G. F. Kuder (1933). Making a rating scale that measures. Personnel Journal, 12:36-40.
`https://psycnet.apa.org/record/1933-04735-001`

Chikio Hayashi(1952). On the prediction of phenomena from qualitative data and the quantification of qualitative data from the mathematical statistical point of view. Annals of the Institute of Statistical Mathematics,3(1): 69-98. `https://ci.nii.ac.jp/naid/30040356125`

Jean-Paul Benzécri (1969). Statistical analysis as a tool to make patterns emerge from data. Methodologies of Pattern Recognition, 35-74.
`https://ci.nii.ac.jp/naid/10006680234`

第二节：联列表

联列表：描述属性变量（定类尺度或定序尺度）的各种状态或相关关系，在调查研究中尤为普遍。

联列表

例：为了了解消费者对自己产品的满意情况，针对不同职业的消费者进行了调查。调查数据整理如下表所示。

评价 职业	非常满意	比较满意	一般	不太满意	不满意	汇总
一般工人 管理者 行政官员 ……..						
汇总						

联列表

例：为了了解消费者对自己产品的满意情况，针对不同职业的消费者进行了调查。调查数据整理如下表所示。

评价 / 职业	非常满意	比较满意	一般	不太满意	不满意	汇总
一般工人 管理者 行政官员 ……						
汇总						

该表可看出：调查对象的职业构成、整体评价、职业分布与评价之间的相关关系。具体地，行列交叉的某个数值越大，表明该类职业与该类评价等级有较强的相关性。

联列表

一般地，假设研究对象有两个属性。其中属性A有r类，分别记为$A_1, A_2, \cdots, A_r$, 属性B有s类，分别记为$B_1, B_2, \cdots, B_s$。属于状态A_i和B_j共同的样本数为$n_{i,j}$, $i = 1, 2, \cdots r, j = 1, 2, \cdots s$。汇总如下表。

	B_1	$\cdots$	B_j	$\ldots$	B_s	合计
A_1	$n_{1,1}$	$\cdots$	$n_{1,j}$	$\cdots$	$n_{1,s}$	$n_{1\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_i	$n_{i,1}$	$\cdots$	$n_{i,j}$	$\cdots$	$n_{i,s}$	$n_{i\cdot}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_r	$n_{r,1}$	$\cdots$	$n_{r,j}$	$\cdots$	$n_{r,s}$	$n_{r\cdot}$
合计	$n_{\cdot 1}$	$\cdots$	$n_{\cdot j}$	$\cdots$	$n_{\cdot s}$	n

联列表

状态A_i的样本数为$n_{i.}=\sum_{j=1}^{s}n_{i,j}$，状态$B_j$的样本数为$n_{.j}=\sum_{i=1}^{r}n_{i,j}$，总样本数为$n=\sum_{i=1}^{r}\sum_{j=1}^{s}n_{i,j}$。为了方便地表示两个属性各个状态之间的关系，往往用频率代替频数，即$p_{i,j}=n_{i,j}/n$。

	B_1	$\cdots$	B_j	$\dots$	B_s	合计
A_1	$p_{1,1}$	$\cdots$	$p_{1,j}$	$\cdots$	$p_{1,s}$	$p_{1.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_i	$p_{i,1}$	$\cdots$	$p_{i,j}$	$\cdots$	$p_{i,s}$	$p_{i.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
A_r	$p_{r,1}$	$\cdots$	$p_{r,j}$	$\cdots$	$p_{r,s}$	$p_{r.}$
合计	$p_{.1}$	$\cdots$	$p_{.j}$	$\cdots$	$p_{.s}$	1

引入矩阵。

$$
\boldsymbol{P}=\begin{pmatrix}
p_{1,1} & \cdots & p_{1,j} & \cdots & p_{1,s}\\
\vdots & \vdots & \vdots & \vdots & \vdots\\
p_{i,1} & \cdots & p_{i,j} & \cdots & p_{i,s}\\
\vdots & \vdots & \vdots & \vdots & \vdots\\
p_{r,1} & \cdots & p_{r,j} & \cdots & p_{r,s}
\end{pmatrix}, \boldsymbol{P_R}=\begin{pmatrix}
p_{1.}\\ \vdots \\ p_{i.} \\ \vdots \\ p_{r.}
\end{pmatrix}, \boldsymbol{P_C}=\begin{pmatrix}
p_{.1}\\ \vdots \\ p_{.j} \\ \vdots \\ p_{.s}
\end{pmatrix}
$$

则有，$\boldsymbol{P}\times \boldsymbol{1}_r=\boldsymbol{P_R}, \boldsymbol{P}^{\mathsf{T}}\times \boldsymbol{1}_s=\boldsymbol{P_C}$.

联列表

概率含义解释。若将$p_{i,j}$视为状态A_i和B_j同时出现的联合概率，则$p_{i.}$表示状态A_i的边缘概率，$p_{.j}$表示属性B_j的边缘概率。

联列表

概率含义解释。若将$p_{i,j}$视为状态A_i和B_j同时出现的联合概率，则$p_{i.}$表示状态A_i的边缘概率，$p_{.j}$表示属性B_j的边缘概率。

考察属性A和属性B之间的相互关系，可以转化为研究各种状态出现的概率。例如，若二者之间相互独立，则下式成立

$$p_{i,j} = p_{i.} \times p_{.j}, i = 1, 2, \cdots, r, j = 1, 2, \cdots, s$$

联列表

概率含义解释。若将$p_{i,j}$视为状态A_i和B_j同时出现的联合概率，则$p_{i.}$表示状态A_i的边缘概率，$p_{.j}$表示属性B_j的边缘概率。

考察属性A和属性B之间的相互关系，可以转化为研究各种状态出现的概率。例如，若二者之间相互独立，则下式成立

$$p_{i,j} = p_{i.} \times p_{.j}, i = 1, 2, \cdots, r, j = 1, 2, \cdots, s$$

由此可得，状态A_i和B_j同时出现的概率的估计

$$\widehat{p}_{i,j} = p_{i.} \times p_{.j}, i = 1, 2, \cdots, r, j = 1, 2, \cdots, s$$

联列表

概率含义解释。若将$p_{i,j}$视为状态A_i和B_j同时出现的联合概率，则$p_{i.}$表示状态A_i的边缘概率，$p_{.j}$表示属性B_j的边缘概率。

考察属性A和属性B之间的相互关系，可以转化为研究各种状态出现的概率。例如，若二者之间相互独立，则下式成立

$$p_{i,j} = p_{i.} \times p_{.j}, i = 1, 2, \cdots, r, j = 1, 2, \cdots, s$$

由此可得，状态A_i和B_j同时出现的概率的估计

$$\widehat{p}_{i,j} = p_{i.} \times p_{.j}, i = 1, 2, \cdots, r, j = 1, 2, \cdots, s$$

实际概率$p_{i,j}$与期望概率$\widehat{p}_{i,j}$的差别大小，可以判断两属性是否独立。

采用假设检验的方法进行量化研究。

1. 提出原假设H_0 :属性A和属性B相互独立；备择假设H_1 :属性A和属性B不独立；
2. 当原假设H_0成立时，检验统计量
   $\chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{[n_{i,j} - n_{i.}n_{.j}/n]^2}{n_{i.}n_{.j}/n} = n \sum_{i=1}^{r} \sum_{j=1}^{s} \frac{[p_{i,j} - p_{i.}p_{.j}]^2}{p_{i.}p_{.j}}$
   近似服从自由度为$(r-1)(s-1)$的χ^2分布；
3. 计算检验统计量的观察值χ^2，如果$\chi^2 > \chi^2_\alpha\big((r-1)(s-1)\big)$，则在显著性水平$\alpha$下拒绝原假设。

联列表

注1：检验统计量刻画了实际频数（频率）与理论频数（频率）之间相对偏差的总加权和，其中权为理论频数（频率）。

注1：检验统计量刻画了实际频数（频率）与理论频数（频率）之间相对偏差的总加权和，其中权为理论频数（频率）。

注2：如果属性变量的取值类别较多，可能导致$p_{i.}$或$p_{.j}$相对较小，从而影响检验统计量的观察值χ^2。

联列表

注1：检验统计量刻画了实际频数（频率）与理论频数（频率）之间相对偏差的总加权和，其中权为理论频数（频率）。

注2：如果属性变量的取值类别较多，可能导致$p_{i.}$或$p_{.j}$相对较小，从而影响检验统计量的观察值χ^2。

注3：联列表可分析属性变量之间的独立性，对于连续性变量按照某种规则分段也适用。

联列表

注1：检验统计量刻画了实际频数（频率）与理论频数（频率）之间相对偏差的总加权和，其中权为理论频数（频率）。

注2：如果属性变量的取值类别较多，可能导致$p_{i.}$或$p_{.j}$相对较小，从而影响检验统计量的观察值χ^2。

注3：联列表可分析属性变量之间的独立性，对于连续性变量按照某种规则分段也适用。

注4：如果拒绝了独立性假设，进一步还希望知道属性变量各状态间的相关关系，则需进行对应分析。

当属性A和B的类别较少时，如果A_i和B_j两个状态同时出现的联合概率$p_{i,j}$相对较大，则说明二者有较强的依赖关系。

当属性A和B的类别较少时，如果A_i和B_j两个状态同时出现的联合概率$p_{i,j}$相对较大，则说明二者有较强的依赖关系。但当属性A和B的类别较多时，很难做出正确判断，此时需要通过降维简化联列表的结构。

第三节：基本理论

当属性A和B的类别较少时，如果A_i和B_j两个状态同时出现的联合概率$p_{i,j}$相对较大，则说明二者有较强的依赖关系。但当属性A和B的类别较多时，很难做出正确判断，此时需要通过降维简化联列表的结构。

对应分析利用降维的思想，在一张二维图上同时表示两类属性变量的各种状态，直观地描述原始数据的结构。

对应分析的数据：频率的联列表。

定义1

当给定状态A_i的条件下，属性B各个状态发生的条件概率构成的向量称之为行剖面，记为$\boldsymbol{P}_i^R = (\frac{p_{i,1}}{p_{i.}}, \cdots, \frac{p_{i,s}}{p_{i.}})^T$。

基本理论

对应分析的数据：频率的联列表。

定义1

当给定状态A_i的条件下，属性B各个状态发生的条件概率构成的向量称之为行剖面，记为$\boldsymbol{P}_i^R = (\frac{p_{i,1}}{p_{i.}}, \cdots, \frac{p_{i,s}}{p_{i.}})^T$。

注：由于$\sum_{j=1}^{s} \frac{p_{i,j}}{p_{i.}} = 1$恒成立。如果将属性$B$各个状态发生的条件概率视为$s$维空间的一个点，那么这些点分布在$s$维空间中的一个面上。

基本理论

基于条件概率，引入距离描述属性不同状态之间的疏远程度。

定义2

属性A的状态A_k和A_l之间的加权欧式距离为

$$D(A_k, A_l)^2 = \sum_{j=1}^{s} \frac{1}{p_{.j}} \left(\frac{p_{k,j}}{p_{k.}} - \frac{p_{l,j}}{p_{l.}}\right)^2 = \sum_{j=1}^{s} \left(\frac{p_{k,j}}{p_{k.}\sqrt{p_{.j}}} - \frac{p_{l,j}}{p_{l.}\sqrt{p_{.j}}}\right)^2$$

基本理论

基于条件概率，引入距离描述属性不同状态之间的疏远程度。

定义2

属性A的状态A_k和A_l之间的加权欧式距离为

$$D(A_k,A_l)^2=\sum_{j=1}^{s}\frac{1}{p_{.j}}\left(\frac{p_{k,j}}{p_{k.}}-\frac{p_{l,j}}{p_{l.}}\right)^2=\sum_{j=1}^{s}\left(\frac{p_{k,j}}{p_{k.}\sqrt{p_{.j}}}-\frac{p_{l,j}}{p_{l.}\sqrt{p_{.j}}}\right)^2$$

注1：式中属性B各个状态的边缘概率的倒数为权。原因在于，当状态B_j出现的边缘概率较大时，则$\left(\frac{p_{k,j}}{p_{k.}}-\frac{p_{l,j}}{p_{l.}}\right)^2$的贡献就被高估了。

基于条件概率，引入距离描述属性不同状态之间的疏远程度。

定义2

属性A的状态A_k和A_l之间的加权欧式距离为

$$D(A_k,A_l)^2=\sum_{j=1}^{s}\frac{1}{p_{.j}}\left(\frac{p_{k,j}}{p_{k.}}-\frac{p_{l,j}}{p_{l.}}\right)^2=\sum_{j=1}^{s}\left(\frac{p_{k,j}}{p_{k.}\sqrt{p_{.j}}}-\frac{p_{l,j}}{p_{l.}\sqrt{p_{.j}}}\right)^2$$

注1：式中属性B各个状态的边缘概率的倒数为权。原因在于，当状态B_j出现的边缘概率较大时，则$\left(\frac{p_{k,j}}{p_{k.}}-\frac{p_{l,j}}{p_{l.}}\right)^2$的贡献就被高估了。

注2：加权距离相当于行剖面的坐标修正为$\left(\frac{p_{i,1}}{p_{i.}\sqrt{p_{.1}}},\cdots,\frac{p_{i,s}}{p_{i.}\sqrt{p_{.s}}}\right)^{\mathrm{T}}$, $i=1,2,\cdots,r$。

若要刻画属性A各种状态的总差异，需引入“重心”概念。

定义3

属性A的重心，为行剖面修正坐标的加权平均

$$\begin{aligned}\overline{\boldsymbol{P}^R} &= \left(\sum_{i=1}^{r} p_{i.}\frac{p_{i,1}}{p_{i.}\sqrt{p_{.1}}}, \cdots, \sum_{i=1}^{r} p_{i.}\frac{p_{i,s}}{p_{i.}\sqrt{p_{.s}}}\right)^T \\ &= \left(\sqrt{p_{.1}}, \cdots, \sqrt{p_{.s}}\right)^T\end{aligned}$$

基本理论

若要刻画属性A各种状态的总差异，需引入“重心”概念。

定义3

属性A的重心，为行剖面修正坐标的加权平均

$$\overline{\boldsymbol{P}^R}=\left(\sum_{i=1}^{r}p_{i.}\frac{p_{i,1}}{p_{i.}\sqrt{p_{.1}}},\cdots,\sum_{i=1}^{r}p_{i.}\frac{p_{i,s}}{p_{i.}\sqrt{p_{.s}}}\right)^T=\left(\sqrt{p_{.1}},\cdots,\sqrt{p_{.s}}\right)^T$$

注：行剖面修正坐标的每个分量，刻画了给定状态A_i时状态B_j发生的条件概率，消除了状态A_i出现的概率影响。重心以此概率为权求平均。

定义4

属性A各状态相对于重心的加权距离总和，即总惯量

$$
\begin{aligned}
\boldsymbol{I}^R &= \sum_{i=1}^{r} p_{i.} D(\boldsymbol{P}_i^R, \overline{\boldsymbol{P}^R})^2 \\
&= \sum_{i=1}^{r}\sum_{j=1}^{s} p_{i.}\left(\frac{p_{i,j}}{p_{i.}\sqrt{p_{.j}}} - \sqrt{p_{.j}}\right)^2 \\
&= \sum_{i=1}^{r}\sum_{j=1}^{s} \frac{(p_{i,j} - p_{i.}p_{.j})^2}{p_{i.}p_{.j}}
\end{aligned}
$$

定义4

属性A各状态相对于重心的加权距离总和，即总惯量

$$
\begin{aligned}
\boldsymbol{I}^R &= \sum_{i=1}^{r} p_{i.} D(\boldsymbol{P}_i^R, \overline{\boldsymbol{P}^R})^2 \\
&= \sum_{i=1}^{r}\sum_{j=1}^{s} p_{i.}\Big(\frac{p_{i,j}}{p_{i.}\sqrt{p_{.j}}} - \sqrt{p_{.j}}\Big)^2 \\
&= \sum_{i=1}^{r}\sum_{j=1}^{s} \frac{(p_{i,j} - p_{i.}p_{.j})^2}{p_{i.}p_{.j}}
\end{aligned}
$$

注：总惯量刻画了属性各状态的差异，且$\boldsymbol{I}^R = \frac{1}{n}\chi^2$。

类似地，也可以按照列进行上述定义。

类似地，也可以按照列进行上述定义。

定义5

当给定状态B_j的条件下，属性A各个状态发生的条件概率构成的向量称之为列剖面，记为$\boldsymbol{P}_j^C=(\frac{p_{1,j}}{p_{\cdot j}},\cdots,\frac{p_{r,j}}{p_{\cdot j}})^T$。

类似地，也可以按照列进行上述定义。

定义5

当给定状态B_j的条件下，属性A各个状态发生的条件概率构成的向量称之为列剖面，记为$\boldsymbol{P}_j^C = (\frac{p_{1,j}}{p_{.j}}, \cdots, \frac{p_{r,j}}{p_{.j}})^T$。

定义6

属性B的状态B_k和B_l之间的加权欧式距离为

$$D(B_k, B_l)^2 = \sum_{i=1}^{r} \frac{1}{p_{i.}} (\frac{p_{i,k}}{p_{.k}} - \frac{p_{i,l}}{p_{.l}})^2 = \sum_{i=1}^{r} (\frac{p_{i,k}}{p_{.k}\sqrt{p_{i.}}} - \frac{p_{i,l}}{p_{.l}\sqrt{p_{i.}}})^2$$

定义7

属性B的重心为

$$\overline{\boldsymbol{P}^C}=\left(\sum_{j=1}^{s}p_{.j}\frac{p_{1,j}}{p_{.j}\sqrt{p_{1.}}},\cdots,\sum_{j=1}^{s}p_{.j}\frac{p_{r,j}}{p_{.j}\sqrt{p_{r.}}}\right)^T=\left(\sqrt{p_{1.}},\cdots,\sqrt{p_{r.}}\right)^T$$

基本理论

定义7

属性B的重心为

$$\overline{\boldsymbol{P}^C} = \Big(\sum_{j=1}^{s} p_{.j}\frac{p_{1,j}}{p_{.j}\sqrt{p_{1.}}}, \cdots, \sum_{j=1}^{s} p_{.j}\frac{p_{r,j}}{p_{.j}\sqrt{p_{r.}}}\Big)^T = \left(\sqrt{p_{1.}}, \cdots, \sqrt{p_{r.}}\right)^T$$

定义8

属性B各状态相对于重心的加权距离总和为

$$\begin{aligned} \boldsymbol{I}^C &= \sum_{i=1}^{r}\sum_{j=1}^{s} p_{.j}\Big(\frac{p_{i,j}}{p_{.j}\sqrt{p_{i.}}} - \sqrt{p_{i.}}\Big)^2 \\ &= \sum_{i=1}^{r}\sum_{j=1}^{s} \frac{(p_{i,j} - p_{i.}p_{.j})^2}{p_{i.}p_{.j}} \end{aligned}$$

基本理论

经过以上分析，对行（列）剖面各点进行变换，计算加权距离来刻画某个属性各个状态之间的接近程度。

基本理论

经过以上分析，对行（列）剖面各点进行变换，计算加权距离来刻画某个属性各个状态之间的接近程度。例如，距离较短的两个状态可化为同一类，反之化为不同的类，分类达到简化数据结构的目的。

基本理论

经过以上分析，对行（列）剖面各点进行变换，计算加权距离来刻画某个属性各个状态之间的接近程度。例如，距离较短的两个状态可化为同一类，反之化为不同的类，分类达到简化数据结构的目的。

该做法不能实现对两个属性同时进行分析。对应分析不求距离而求协方差，然后提取主因子，对两个属性的各状态进行分析。

基本理论

对行剖面进行分析。

为消除属性B各状态发生的边缘概率的影响，对每个行剖面的坐标修正，相当于进行变换

$$
\begin{aligned}
\left(\frac{p_{i,1}}{p_{i.}\sqrt{p_{.1}}}, \cdots, \frac{p_{i,s}}{p_{i.}\sqrt{p_{.s}}}\right) &= \left(\left(\frac{p_{i,1}}{p_{i.}}, \cdots, \frac{p_{i,s}}{p_{i.}}\right) diag(\sqrt{p_{.1}}, \cdots, \sqrt{p_{.s}}\right) \\
&\triangleq (\boldsymbol{P}_i^R)^{\mathsf{T}} \Phi^R
\end{aligned}
$$

其中$\Phi^R = diag(\sqrt{p_{.1}}, \sqrt{p_{.2}}, \cdots, \sqrt{p_{.s}})$。

基本理论

对行剖面进行分析。

为消除属性B各状态发生的边缘概率的影响，对每个行剖面的坐标修正，相当于进行变换

$$
\begin{aligned}
\left(\frac{p_{i,1}}{p_{i.}\sqrt{p_{.1}}}, \cdots, \frac{p_{i,s}}{p_{i.}\sqrt{p_{.s}}}\right) &= \left(\left(\frac{p_{i,1}}{p_{i.}}, \cdots, \frac{p_{i,s}}{p_{i.}}\right) diag(\sqrt{p_{.1}}, \cdots, \sqrt{p_{.s}}\right) \\
&\triangleq (\boldsymbol{P}_i^R)^{\mathsf{T}} \Phi^R
\end{aligned}
$$

其中$\Phi^R = diag(\sqrt{p_{.1}}, \sqrt{p_{.2}}, \cdots, \sqrt{p_{.s}})$。从而行剖面构成一个新的矩阵

$$
\begin{pmatrix}
(\boldsymbol{P}_1^R)^{\mathsf{T}} \Phi^R \\
\vdots \\
(\boldsymbol{P}_r^R)^{\mathsf{T}} \Phi^R
\end{pmatrix}
$$

基本理论

对应的加权协差阵为s阶方阵,

$$\Sigma^R = \sum_{i=1}^{r} p_{i.} \left[(\boldsymbol{P}_i^R)^{\mathsf{T}} \Phi^R - (\overline{\boldsymbol{P}^R})^{\mathsf{T}} \right]^{\mathsf{T}} \left[(\boldsymbol{P}_i^R)^{\mathsf{T}} \Phi^R - (\overline{\boldsymbol{P}^R})^{\mathsf{T}} \right]$$

基本理论

对应的加权协差阵为s阶方阵，

$$\Sigma^R = \sum_{i=1}^{r} p_{i.}\left[(\boldsymbol{P}_i^R)^\mathsf{T}\Phi^R - (\overline{\boldsymbol{P}^R})^\mathsf{T}\right]^\mathsf{T}\left[(\boldsymbol{P}_i^R)^\mathsf{T}\Phi^R - (\overline{\boldsymbol{P}^R})^\mathsf{T}\right]$$

同时，对于任意的$i = 1, 2, \cdots, r$满足

$$\begin{aligned}
&\left[(\boldsymbol{P}_i^R)^\mathsf{T}\Phi^R - (\overline{\boldsymbol{P}^R})^\mathsf{T}\right]\overline{\boldsymbol{P}^R} \\
=&\left[\frac{p_{i,1}}{p_{i.}\sqrt{p_{.1}}} - \sqrt{p_{.1}}, \cdots, \frac{p_{i,s}}{p_{i.}\sqrt{p_{.s}}} - \sqrt{p_{.s}}\right]\left[\sqrt{p_{.1}}, \cdots, \sqrt{p_{.s}}\right]^\mathsf{T} \\
=&\left[\frac{p_{i,1} - p_{i.}p_{.1}}{p_{i.}\sqrt{p_{.1}}}, \cdots, \frac{p_{i,s} - p_{i.}p_{.s}}{p_{i.}\sqrt{p_{.s}}}\right]\left[\sqrt{p_{.1}}, \cdots, \sqrt{p_{.s}}\right]^\mathsf{T} \\
=&0
\end{aligned}$$

基本理论

对应的加权协差阵为s阶方阵，

$$\Sigma^R = \sum_{i=1}^{r} p_{i.}\big[(\boldsymbol{P}_i^R)^\mathsf{T}\Phi^R - (\overline{\boldsymbol{P}^R})^\mathsf{T}\big]^\mathsf{T}\big[(\boldsymbol{P}_i^R)^\mathsf{T}\Phi^R - (\overline{\boldsymbol{P}^R})^\mathsf{T}\big]$$

同时，对于任意的$i = 1, 2, \cdots, r$满足

$$\begin{aligned}
&\big[(\boldsymbol{P}_i^R)^\mathsf{T}\Phi^R - (\overline{\boldsymbol{P}^R})^\mathsf{T}\big]\overline{\boldsymbol{P}^R} \\
=&\Big[\frac{p_{i,1}}{p_{i.}\sqrt{p_{.1}}} - \sqrt{p_{.1}}, \cdots, \frac{p_{i,s}}{p_{i.}\sqrt{p_{.s}}} - \sqrt{p_{.s}}\Big]\big[\sqrt{p_{.1}}, \cdots, \sqrt{p_{.s}}\big]^\mathsf{T} \\
=&\Big[\frac{p_{i,1} - p_{i.}p_{.1}}{p_{i.}\sqrt{p_{.1}}}, \cdots, \frac{p_{i,s} - p_{i.}p_{.s}}{p_{i.}\sqrt{p_{.s}}}\Big]\big[\sqrt{p_{.1}}, \cdots, \sqrt{p_{.s}}\big]^\mathsf{T} \\
=&0
\end{aligned}$$

因而有$\Sigma^R\overline{\boldsymbol{P}^R} = 0 \times \overline{\boldsymbol{P}^R} = \mathbf{0}$,

对应的加权协差阵为s阶方阵，

$$\Sigma^R=\sum_{i=1}^{r}p_{i.}\left[(\boldsymbol{P}_i^R)^{\mathsf{T}}\Phi^R-(\overline{\boldsymbol{P}^R})^{\mathsf{T}}\right]^{\mathsf{T}}\left[(\boldsymbol{P}_i^R)^{\mathsf{T}}\Phi^R-(\overline{\boldsymbol{P}^R})^{\mathsf{T}}\right]$$

同时，对于任意的$i=1,2,\cdots,r$满足

$$\begin{aligned}&\left[(\boldsymbol{P}_i^R)^{\mathsf{T}}\Phi^R-(\overline{\boldsymbol{P}^R})^{\mathsf{T}}\right]\overline{\boldsymbol{P}^R}\\=&\left[\frac{p_{i,1}}{p_{i.}\sqrt{p_{.1}}}-\sqrt{p_{.1}},\cdots,\frac{p_{i,s}}{p_{i.}\sqrt{p_{.s}}}-\sqrt{p_{.s}}\right]\left[\sqrt{p_{.1}},\cdots,\sqrt{p_{.s}}\right]^{\mathsf{T}}\\=&\left[\frac{p_{i,1}-p_{i.}p_{.1}}{p_{i.}\sqrt{p_{.1}}},\cdots,\frac{p_{i,s}-p_{i.}p_{.s}}{p_{i.}\sqrt{p_{.s}}}\right]\left[\sqrt{p_{.1}},\cdots,\sqrt{p_{.s}}\right]^{\mathsf{T}}\\=&0\end{aligned}$$

因而有$\Sigma^R\overline{\boldsymbol{P}^R}=0\times\overline{\boldsymbol{P}^R}=\mathbf{0}$, 意味着重心$\overline{\boldsymbol{P}^R}$是协差阵的特征值为0所对应的特征向量。

基本理论

$\mathbf{\Sigma}^R$的代表元素为

$$\begin{aligned} a_{i,j} &= \sum_{k=1}^{r} p_{k.}\Big(\frac{p_{k,i}}{p_{k.}\sqrt{p_{.i}}} - \sqrt{p_{.i}}\Big)\Big(\frac{p_{k,j}}{p_{k.}\sqrt{p_{.j}}} - \sqrt{p_{.j}}\Big) \\ &= \sum_{k=1}^{r} \Big(\frac{p_{k,i} - p_{k.}p_{.i}}{\sqrt{p_{k.}p_{.i}}}\Big)\Big(\frac{p_{k,j} - p_{k.}p_{.j}}{\sqrt{p_{k.}p_{.j}}}\Big) \\ &\triangleq \sum_{k=1}^{r} z_{k,i} z_{k,j} \end{aligned}$$

其中$z_{i,j} = \frac{p_{i,j} - p_{i.}p_{.j}}{\sqrt{p_{i.}p_{.j}}}, i = 1, \cdots, r; j = 1, \cdots, s$。若令$Z = (z_{i,j})_{r\times s}$，则有$\mathbf{\Sigma}^R = Z^{\mathsf{T}}Z$.

类似地，也可以对列剖面进行分析。加权协差阵为r阶方阵

$$\boldsymbol{\Sigma}^C = \sum_{j=1}^{s} p_{\cdot j}\left[(\boldsymbol{P}_j^C \Phi^C)^\mathsf{T} - (\overline{\boldsymbol{P}^C})^\mathsf{T}\right]^\mathsf{T}\left[(\boldsymbol{P}_j^C \Phi^C)^\mathsf{T} - (\overline{\boldsymbol{P}^C})^\mathsf{T}\right]$$

其中$\Phi^C = diag(\sqrt{p_{1\cdot}}, \cdots, \sqrt{p_{r\cdot}})$。满足

- $\boldsymbol{\Sigma}^C = ZZ^\mathsf{T}$;
- 重心$\overline{\boldsymbol{P}^C} = (\sqrt{p_{1\cdot}}, \cdots, \sqrt{p_{r\cdot}})^\mathsf{T}$ 是其特征值为0所对应的特征向量。

基本理论

根据矩阵知识有，s阶方阵$\mathbf{\Sigma}^R = Z^\top Z$与$r$阶方阵$\mathbf{\Sigma}^C = ZZ^\top$

- 有相同的特征根;
- 都有一个零根, 非零特征根最多为$\min(r, s) - 1$.

基本理论

根据矩阵知识有，s阶方阵$\boldsymbol{\Sigma}^R = Z^\mathsf{T}Z$与$r$阶方阵$\boldsymbol{\Sigma}^C = ZZ^\mathsf{T}$

- 有相同的特征根;
- 都有一个零根, 非零特征根最多为$\min(r, s) - 1$.

不妨记$\boldsymbol{\Sigma}^R$的非零特征根为$\lambda_1 \geq \cdots \geq \lambda_k$，其中$0 < k \leq \min(r, s) - 1$，对应的特征向量分别为$\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_k$，即

$$Z^\mathsf{T}Z\boldsymbol{\xi}_i = \boldsymbol{\Sigma}^R\boldsymbol{\xi}_i = \lambda_i\boldsymbol{\xi}_i, i = 1, 2, \cdots, k$$

基本理论

根据矩阵知识有，s阶方阵$\mathbf{\Sigma}^R = Z^\mathsf{T} Z$与$r$阶方阵$\mathbf{\Sigma}^C = ZZ^\mathsf{T}$

- 有相同的特征根;
- 都有一个零根, 非零特征根最多为$\min(r, s) - 1$.

不妨记$\mathbf{\Sigma}^R$的非零特征根为$\lambda_1 \geq \cdots \geq \lambda_k$，其中$0 < k \leq \min(r, s) - 1$，对应的特征向量分别为$\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_k$，即

$$Z^\mathsf{T} Z \boldsymbol{\xi}_i = \mathbf{\Sigma}^R \boldsymbol{\xi}_i = \lambda_i \boldsymbol{\xi}_i, i = 1, 2, \cdots, k$$

上式两边左乘Z有，

$$ZZ^\mathsf{T} Z \boldsymbol{\xi}_i = \mathbf{\Sigma}^C (Z \boldsymbol{\xi}_i) = \lambda_i (Z \boldsymbol{\xi}_i)$$

根据矩阵知识有，s阶方阵$\mathbf{\Sigma}^R = Z^{\mathsf{T}}Z$与$r$阶方阵$\mathbf{\Sigma}^C = ZZ^{\mathsf{T}}$

- 有相同的特征根;
- 都有一个零根, 非零特征根最多为$\min(r, s) - 1$.

不妨记$\mathbf{\Sigma}^R$的非零特征根为$\lambda_1 \geq \cdots \geq \lambda_k$，其中$0 < k \leq \min(r, s) - 1$，对应的特征向量分别为$\boldsymbol{\xi}_1, \cdots, \boldsymbol{\xi}_k$，即

$$Z^{\mathsf{T}}Z\boldsymbol{\xi}_i = \mathbf{\Sigma}^R\boldsymbol{\xi}_i = \lambda_i\boldsymbol{\xi}_i, i = 1, 2, \cdots, k$$

上式两边左乘Z有，

$$ZZ^{\mathsf{T}}Z\boldsymbol{\xi}_i = \mathbf{\Sigma}^C(Z\boldsymbol{\xi}_i) = \lambda_i(Z\boldsymbol{\xi}_i)$$

即$Z\boldsymbol{\xi}_i$为矩阵$\mathbf{\Sigma}^C$的特征向量。

注1：对联列表数据做Q型因子分析（对属性A进行），可方便地得到R型因子分析（对属性B进行）的结果。

基本理论

注1：对联列表数据做Q型因子分析（对属性A进行），可方便地得到R型因子分析（对属性B进行）的结果。

注2：原始变量的总方差为行（列）剖面的总惯量；提取公共因子的总方差为特征根之和，显然有$\boldsymbol{I}^R = \boldsymbol{I}^C = \sum_{i=1}^{k} \lambda_i$。

基本理论

注1：对联列表数据做Q型因子分析（对属性A进行），可方便地得到R型因子分析（对属性B进行）的结果。

注2：原始变量的总方差为行（列）剖面的总惯量；提取公共因子的总方差为特征根之和，显然有$\boldsymbol{I}^R = \boldsymbol{I}^C = \sum_{i=1}^{k} \lambda_i$。

注3：相同数量的公共因子，分别解释两个属性的方差比例相同。一般取2个公共因子，则可在二维图上同时画出两个属性的各个状态。

基本理论

注1：对联列表数据做Q型因子分析（对属性A进行），可方便地得到R型因子分析（对属性B进行）的结果。

注2：原始变量的总方差为行（列）剖面的总惯量；提取公共因子的总方差为特征根之和，显然有$\boldsymbol{I}^R=\boldsymbol{I}^C=\sum_{i=1}^k\lambda_i$。

注3：相同数量的公共因子，分别解释两个属性的方差比例相同。一般取2个公共因子，则可在二维图上同时画出两个属性的各个状态。

注4：对应分析只能直观展示属性之间的关联关系，不能给出关联关系的度量刻画。

对应分析步骤

1. 原始数据整理汇总得到频数的联列表；
2. 计算频率联列表；
3. 计算Z矩阵；
4. 由$\mathbf{\Sigma}^R$（或$\mathbf{\Sigma}^C$）对属性A进行Q型（或R型）因子分析，由此导出R型（或Q型）因子分析的结果；
5. 在二维图上同时画出属性A和B各个状态。