Factor Analysis

因子分析

肖 磊，2026 年 5 月 12 日

大纲

因子分析 (Factor Analysis)

- 正交因子模型 (The Orthogonal Factor Model)
- 因子模型的估计 (Estimation of the Factor Model)
- 因子得分与策略 (Factor Scores and Strategies)
- 波士顿住房数据 (Boston Housing)
- 能力与智力测试 (Ability and Intelligence Tests)

前言 (Preface)

- 将 (多元变量表示的) 相关信息看作来自有限数量的潜在因子进行建模.
  - 例如，一项关于家庭消费的调查中，观察到了一个月内 p 种不同商品的消费水平 $\boldsymbol{X}$.
  - 在整个调查过程中，$\boldsymbol{X}$ 的 p 个组成部分的变化和协变，实际上可能由家庭的两到三个主要社会行为因子 (factors) 来解释.

    对舒适的基本需求
    追求一定社会地位的意愿
    其他社会潜在因素
  - 对于社会学家而言，这些无法直接观测到的因素，远比能够直接观测到的量化指标 ($\boldsymbol{X}$) 本身更具吸引力.
  - 它们能让人们更好地理解家庭的行为表现.
  - 这一类因子分析 (factor analysis) 在心理学、市场营销、经济学、政治学等诸多领域备受关注.

前言 (Preface)

- 将 (多元变量表示的) 相关信息看作来自有限数量的潜在因子进行建模.
  - 因子分析的目的:

    ① 如何提供一个解决这些问题的统计模型?

    ② 如何解释所得到的模型?
  - 本章的核心统计主旨同样是降低观测数据的维度.
  - 然而，这里所采用的视角有所不同：我们假定存在一种模型 (称为因子模型)，该模型表明，X 的 p 个元素之间的大部分协方差，可由数量有限的潜在因子来解释.

正交因子模型 (The Orthogonal Factor Model)

- 因子分析 (factor analysis) 的目的是使用较少的变量 (即所谓的因子) 来解释数据矩阵 $\mathscr{X}$ 中 p 个变量的结果.
  - 理想情况下，数据矩阵 $\mathscr{X}$ 中的所有信息都可以由数量更少的因子 (factors) 重现.
  - 这些因子可解释为所观测到的 $\boldsymbol{x} \in \mathbb{R}^p$ 的潜在的 (无法直接观察) 共同特征.

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix}$$

因子，$\ell = 1, 2, \ldots, k$, 且 $k \lll p$

$$x_j = \sum_{\ell=1}^{k} \left(q_{j\ell} f_\ell \right) + \mu_j \,, \quad j = 1, 2, \ldots, p$$

$$\boldsymbol{X} = \mathscr{Q} \boldsymbol{F} + \boldsymbol{\mu}$$

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} = \begin{pmatrix} q_{11}f_1 + q_{12}f_2 + \cdots + q_{1k}f_k + \mu_1 \\ q_{21}f_1 + q_{22}f_2 + \cdots + q_{2k}f_k + \mu_2 \\ \vdots \\ q_{p1}f_1 + q_{p2}f_2 + \cdots + q_{pk}f_k + \mu_p \end{pmatrix} = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1k} \\ q_{21} & q_{22} & \cdots & q_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ q_{p1} & q_{p2} & \cdots & q_{pk} \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_k \end{pmatrix} + \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix}$$

正交因子模型 (The Orthogonal Factor Model)

- 因子分析 (factor analysis) 的目的是使用较少的变量 (即所谓的因子) 来解释数据矩阵 $\mathscr{X}$ 中 p 个变量的结果.
  - 模型的矩阵形式:

$$\boldsymbol{X} = \mathcal{Q}\,\boldsymbol{F} + \boldsymbol{\mu}$$

  - 在使用因子模型时，通常假定因子 $\boldsymbol{F}$ 是中心化、不相关且经过标准化处理的:

$$\mathrm{E}(\boldsymbol{F}) = \boldsymbol{0}\,,\quad \mathrm{Var}(\boldsymbol{F}) = \mathscr{I}_k$$

正交因子模型 (The Orthogonal Factor Model)

- 事实上，通过主成分 (principal components) 来创建观测数据的一种表征是可行的.
  - 设 $\mathrm{E}(\boldsymbol{X}) = \boldsymbol{\mu}$， $\mathrm{Var}(\boldsymbol{X}) = \Sigma$.
  - 如果 Σ 的后 $p-k$ 个特征值为零，我们可以将 $\boldsymbol{X}$ 表示为因子模型 $\boldsymbol{X} = \mathcal{Q}\,\boldsymbol{F} + \boldsymbol{\mu}$.
  - Σ 的谱分解为 $\Sigma = \Gamma \Lambda \Gamma^{\mathrm{T}}$.
  - 假设 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0$ 且 $\lambda_{k+1} = \cdots = \lambda_p = 0$.

$$\Longrightarrow \quad \Sigma = \Gamma \Lambda \Gamma^{\mathrm{T}} = \begin{pmatrix} \Gamma_1 & \Gamma_2 \end{pmatrix} \begin{pmatrix} \Lambda_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \Gamma_1^{\mathrm{T}} \\ \Gamma_2^{\mathrm{T}} \end{pmatrix} = \sum_{\ell=1}^{k} \lambda_\ell \boldsymbol{\gamma}_\ell \boldsymbol{\gamma}_\ell^{\mathrm{T}}$$

正交因子模型 (The Orthogonal Factor Model)

- 事实上，通过主成分 (principal components) 来创建观测数据的一种表征是可行的.
  - 主成分由 $\boldsymbol{Y} = \Gamma^{\mathrm{T}}(\boldsymbol{X}-\boldsymbol{\mu})$ 定义.

$$\Longrightarrow \quad \Gamma \boldsymbol{Y} = \underbrace{\Gamma\Gamma^{\mathrm{T}}}_{\mathcal{I}}(\boldsymbol{X}-\boldsymbol{\mu}) = \boldsymbol{X}-\boldsymbol{\mu}$$

$$\Longrightarrow \quad \boldsymbol{Y} = \begin{pmatrix} \boldsymbol{Y}_1 \\ \boldsymbol{Y}_2 \end{pmatrix} = \Gamma^{\mathrm{T}}(\boldsymbol{X}-\boldsymbol{\mu}) = \begin{pmatrix} \Gamma_1^{\mathrm{T}} \\ \Gamma_2^{\mathrm{T}} \end{pmatrix}(\boldsymbol{X}-\boldsymbol{\mu}) \sim \left(\boldsymbol{0}\,,\ \begin{pmatrix} \Lambda_1 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}\right)$$

$\Longrightarrow$ $\boldsymbol{Y}_2$ 具有奇异分布，其均值和协方差矩阵均为零.

$$\Longrightarrow \quad \boldsymbol{X}-\boldsymbol{\mu} = \Gamma\boldsymbol{Y} = (\Gamma_1 \quad \Gamma_2)\begin{pmatrix} \boldsymbol{Y}_1 \\ \boldsymbol{Y}_2 \end{pmatrix} = \Gamma_1\boldsymbol{Y}_1 + \Gamma_2\boldsymbol{Y}_2 = \Gamma_1\boldsymbol{Y}_1$$

$$\Longrightarrow \quad \boldsymbol{X} = \Gamma_1\boldsymbol{Y}_1 + \boldsymbol{\mu} = \underbrace{\Gamma_1\,\Lambda^{1/2}}_{\mathcal{Q}}\ \underbrace{\Lambda^{-1/2}\,\boldsymbol{Y}_1}_{\boldsymbol{F}} + \boldsymbol{\mu} \quad \Longrightarrow \quad \boldsymbol{X} = \mathcal{Q}\boldsymbol{F} + \boldsymbol{\mu}$$

$$\mathrm{E}(\boldsymbol{F}) = \boldsymbol{0}\ , \quad \mathrm{Var}(\boldsymbol{F}) = \mathcal{I}_k$$

正交因子模型 (The Orthogonal Factor Model)

- 对模型: $\boldsymbol{X} = \mathcal{Q}\,\boldsymbol{F} + \boldsymbol{\mu}$, 注意到

$$
\begin{aligned}
\Sigma &= \mathrm{E}\left[(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})^{\mathrm{T}}\right] \\
&= \mathrm{E}\left[(\mathcal{Q}\boldsymbol{F}+\boldsymbol{\mu}-\boldsymbol{\mu})(\mathcal{Q}\boldsymbol{F}+\boldsymbol{\mu}-\boldsymbol{\mu})^{\mathrm{T}}\right] \\
&= \mathcal{Q}\,\mathrm{E}\left(\boldsymbol{F}\boldsymbol{F}^{\mathrm{T}}\right)\mathcal{Q}^{\mathrm{T}} \qquad \mathrm{Var}(\boldsymbol{F}) = \mathcal{I}_k \\
&= \mathcal{Q}\mathcal{Q}^{\mathrm{T}} = \sum_{j=1}^{k} \lambda_j \boldsymbol{\gamma}_j \boldsymbol{\gamma}_j^{\mathrm{T}}
\end{aligned}
$$

  - 这说明变量 $\boldsymbol{X}$ 可以由 k 个 $(k < p)$ 不相关因子的加权和完全确定.
  - 然而，推导过程中所采用的情形过于理想化.
  - 在实际中，协方差矩阵很少是奇异的.

正交因子模型 (The Orthogonal Factor Model)

- 在因子分析 (factor analysis) 中，常见做法是将因子的影响分为公共影响和特殊影响.
  - 存在对 $\boldsymbol{X}$ 所有组成部分都通用的、信息量大的因子.
  - 也存在一些只特定作用于 $\boldsymbol{X}$ 某些组成部分的因子.
  - 实践中使用的因子分析模型是

$$\boldsymbol{X} = \mathcal{Q}\,\boldsymbol{F} + \boldsymbol{U} + \boldsymbol{\mu}$$

$\mathcal{Q}_{p\times k}$ 公因子载荷

$\boldsymbol{F}_{k\times 1}$ 公共因子

$\boldsymbol{U}_{p\times 1}$ 特殊因子

  - 假设:

$$\begin{aligned}
\mathrm{E}(\boldsymbol{F}) &= \boldsymbol{0}\\
\mathrm{Var}(\boldsymbol{F}) &= \mathcal{I}_k\\
\mathrm{E}(\boldsymbol{U}) &= \boldsymbol{0}\\
\mathrm{Cov}\left(U_i, U_j\right) &= 0, \quad i \neq j\\
\mathrm{Cov}(\boldsymbol{F}, \boldsymbol{U}) &= \boldsymbol{0}
\end{aligned}$$

$$\mathrm{Var}(\boldsymbol{U}) \triangleq \boldsymbol{\Psi} = \begin{pmatrix} \psi_{11} & & & \\ & \psi_{22} & & \\ & & \ddots & \\ & & & \psi_{pp} \end{pmatrix}$$

正交因子模型 (The Orthogonal Factor Model)

正交因子模型

$$\underset{(p\times 1)}{\boldsymbol{X}} = \underset{(p\times k)}{\mathcal{Q}}\ \underset{(k\times 1)}{\boldsymbol{F}} + \underset{(p\times 1)}{\boldsymbol{U}} + \underset{(p\times 1)}{\boldsymbol{\mu}}$$

μ_j = 第 j 个变量的均值

U_j = 第 j 个特殊因子

F_ℓ = 第 ℓ 个公共因子

$q_{j\ell}$ = 第 j 个变量在第 ℓ 个公共因子上的载荷

随机向量 $\boldsymbol{F}$ 与 $\boldsymbol{U}$ 是不可观测且不相关的.

正交因子模型 (The Orthogonal Factor Model)

$$\begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_p \end{pmatrix} = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1k} \\ q_{21} & q_{22} & \cdots & q_{2k} \\ \vdots & \vdots & & \vdots \\ q_{p1} & q_{p2} & \cdots & q_{pk} \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \\ \vdots \\ F_k \end{pmatrix} + \begin{pmatrix} U_1 \\ U_2 \\ \vdots \\ U_p \end{pmatrix} + \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_p \end{pmatrix}$$

$$\Longrightarrow \quad X_j = q_{j1}F_1 + q_{j2}F_2 + \cdots + q_{jk}F_k + U_j + \mu_j$$

$$= \sum_{\ell=1}^{k} q_{j\ell}F_\ell + U_j + \mu_j\,, \quad j = 1,\ 2,\ \ldots,\ p$$

$$\mathrm{Var}(\boldsymbol{F}) = \mathcal{I}_k$$

$$\Longrightarrow \quad \sigma_{X_jX_j} = \mathrm{Var}\left(X_j\right) = \sum_{\ell=1}^{k} q_{j\ell}^2\, \mathrm{Var}\left(F_\ell\right) + \mathrm{Var}\left(U_j\right)$$

$$= \sum_{\ell=1}^{k} q_{j\ell}^2 + \psi_{jj}$$

$$\mathrm{Var}(\boldsymbol{U}) \triangleq \boldsymbol{\Psi} = \begin{pmatrix} \psi_{11} & & & \\ & \psi_{22} & & \\ & & \ddots & \\ & & & \psi_{pp} \end{pmatrix}$$

公因子方差: $h_j^2 = \sum_{\ell=1}^{k} q_{j\ell}^2$

特殊因子方差: ψ_{jj}

正交因子模型 (The Orthogonal Factor Model)

$$\begin{aligned}\mathrm{Var}(\boldsymbol{X}) = \Sigma = \mathrm{E}\left[(\boldsymbol{X}-\boldsymbol{\mu})(\boldsymbol{X}-\boldsymbol{\mu})^{\mathrm{T}}\right] &= \mathrm{E}\left[(\mathcal{Q}\boldsymbol{F}+\boldsymbol{U})(\mathcal{Q}\boldsymbol{F}+\boldsymbol{U})^{\mathrm{T}}\right] \\ &= \mathrm{E}\left[(\mathcal{Q}\boldsymbol{F}+\boldsymbol{U})\left(\boldsymbol{F}^{\mathrm{T}}\mathcal{Q}^{\mathrm{T}}+\boldsymbol{U}^{\mathrm{T}}\right)\right] = \mathrm{E}\left(\mathcal{Q}\boldsymbol{F}\boldsymbol{F}^{\mathrm{T}}\mathcal{Q}^{\mathrm{T}}+\mathcal{Q}\boldsymbol{F}\boldsymbol{U}^{\mathrm{T}}+\boldsymbol{U}\boldsymbol{F}^{\mathrm{T}}\mathcal{Q}^{\mathrm{T}}+\boldsymbol{U}\boldsymbol{U}^{\mathrm{T}}\right) \\ &= \mathcal{Q}\,\mathrm{E}\left(\boldsymbol{F}\boldsymbol{F}^{\mathrm{T}}\right)\mathcal{Q}^{\mathrm{T}}+\mathrm{E}\left(\boldsymbol{U}\boldsymbol{U}^{\mathrm{T}}\right) = \mathcal{Q}\,\mathrm{Var}(\boldsymbol{F})\,\mathcal{Q}^{\mathrm{T}}+\mathrm{Var}(\boldsymbol{U}) = \mathcal{Q}\mathcal{Q}^{\mathrm{T}}+\varPsi\end{aligned}$$

- 从某种意义上说，因子模型在很大程度上通过少量对其 p 个分量通用的潜在因子 $\boldsymbol{F}$ 来解释 $\boldsymbol{X}$ 的变化，并完整地解释其各分量之间的所有相关结构，再加上一些 “噪声” $\boldsymbol{U}$，它使得每个分量的特定变化得以纳入.
- 调整特殊因子以反映每个成分的个体差异.
- 因子分析 (Factor analysis) 依赖于上述假设.
- 如果这些假设不成立，分析结果可能是荒谬的.

正交因子模型 (The Orthogonal Factor Model)

- 虽然主成分分析和因子分析可能存在关联，但它们在本质上有很大差异.
  - 主成分是对 $\boldsymbol{X}$ 进行的线性变换，按方差从大到小排列，用于降低数据集的维度.
  - 在因子分析中，通过有限数量的潜在因子的线性变换，对 $\boldsymbol{X}$ 的变化进行建模.
- 因子分析的目标是找到载荷矩阵 $\mathcal{Q}$ 和特殊方差矩阵 Ψ.
  - $\mathcal{Q}$ 和 Ψ 的估计是从协方差结构 $\Sigma = \mathcal{Q}\mathcal{Q}^{\mathrm{T}} + \Psi$ 推导得出的.

正交因子模型 (The Orthogonal Factor Model)

- 因子的解释
  - 假设包含 k 个因子的因子模型是合理的，也就是说，$\boldsymbol{X}$ 中 p 个变量的大部分变化可以由这 k 个固定的潜在因子来解释.
  - 接下来顺理成章的一步，就是要弄清楚这些因素代表着什么.
  - 为了解释 F_ℓ，先计算它与原始变量 X_j 的相关性是有意义的.

$$\rho_{F_\ell X_j}, \quad \ell = 1, 2, \ldots, k\,; \; j = 1, 2, \ldots, p$$

正交因子模型 (The Orthogonal Factor Model)

- 因子的解释

$$\Sigma_{XF} = \mathrm{Var}(\boldsymbol{X}, \boldsymbol{F}) = \mathrm{E}\left[(\boldsymbol{X}-\boldsymbol{\mu})\boldsymbol{F}^{\mathrm{T}}\right] = \mathrm{E}\left[(\mathcal{Q}\boldsymbol{F}+\boldsymbol{U})\boldsymbol{F}^{\mathrm{T}}\right]$$

$$\boldsymbol{X} = \mathcal{Q}\boldsymbol{F} + \boldsymbol{U} + \boldsymbol{\mu}$$

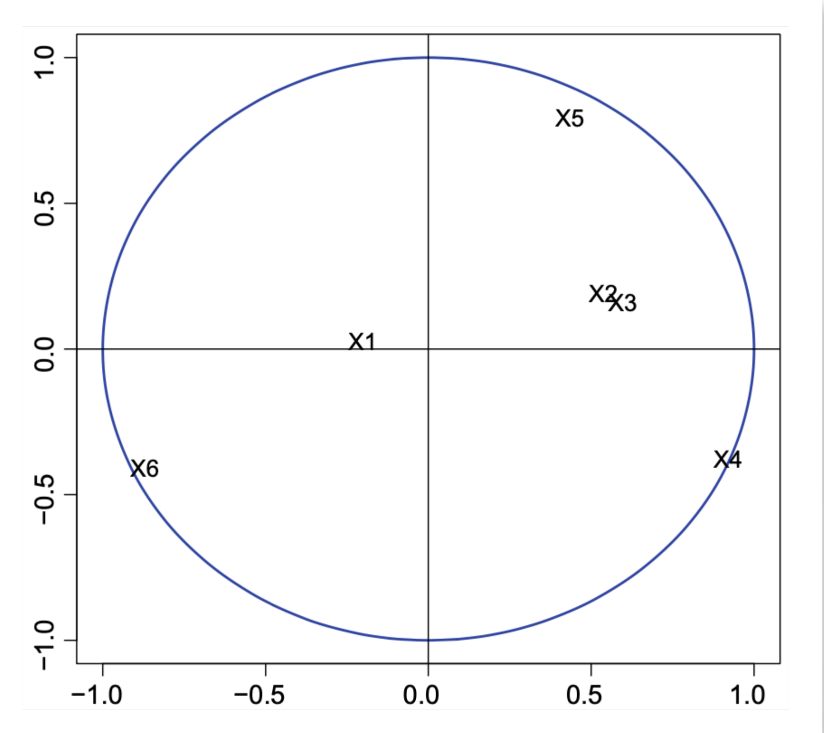
$$= \mathrm{E}\left(\mathcal{Q}\boldsymbol{F}\boldsymbol{F}^{\mathrm{T}} + \boldsymbol{U}\boldsymbol{F}^{\mathrm{T}}\right) = \mathcal{Q}\,\mathrm{E}\left(\boldsymbol{F}\boldsymbol{F}^{\mathrm{T}}\right) + \mathrm{E}\left(\boldsymbol{U}\boldsymbol{F}^{\mathrm{T}}\right) = \mathcal{Q}$$

$\Longrightarrow$ 相关矩阵为 $P_{XF} = D^{-1/2}\mathcal{Q}$

$$\rho_{F_\ell X_j}$$

$$D = \begin{pmatrix} \sigma_{X_1X_1} & & & \\ & \sigma_{X_2X_2} & & \\ & & \ddots & \\ & & & \sigma_{X_pX_p} \end{pmatrix}$$

- 与主成分分析 (PCA) 类似，可以作一个图形来分析原始变量 $X_1, X_2, \ldots, X_p$ 中哪些在不可观测的公共因子 $F_1, F_2, \ldots, F_k$ 中发挥作用.



正交因子模型 (The Orthogonal Factor Model)

- 尺度不变性

  - 如果将 $\boldsymbol{X}$ 的度量尺度变为 $\boldsymbol{Y}=\mathscr{C}\boldsymbol{X}$ 会发生什么情况呢？其中

$$\mathscr{C}=\begin{pmatrix} c_1 & & & \\ & c_2 & & \\ & & \ddots & \\ & & & c_p \end{pmatrix}$$

  - 如果对于 $\boldsymbol{X}$，含有 k 个因子的因子模型成立，其中 $\mathcal{Q}=\mathcal{Q}_X$，$\Psi=\Psi_X$，即

$$X_j=\sum_{\ell=1}^{k} q_{j\ell}F_\ell+U_j+\mu_j\,,\quad j=1,\,2,\,\ldots,\,p$$

$$\Sigma=\mathcal{Q}_X\mathcal{Q}_X^{\mathrm{T}}+\Psi_X$$

$$\Longrightarrow\quad \mathrm{Var}(\boldsymbol{Y})=\mathrm{Var}(\mathscr{C}\boldsymbol{X})=\mathscr{C}\ \mathrm{Var}(\boldsymbol{X})\ \mathscr{C}^{\mathrm{T}}=\mathscr{C}\Sigma\mathscr{C}^{\mathrm{T}}=\mathscr{C}\left(\mathcal{Q}_X\mathcal{Q}_X^{\mathrm{T}}+\Psi_X\right)\mathscr{C}^{\mathrm{T}}$$

$$=\mathscr{C}\mathcal{Q}_X\mathcal{Q}_X^{\mathrm{T}}\mathscr{C}^{\mathrm{T}}+\mathscr{C}\Psi_X\mathscr{C}^{\mathrm{T}}=\mathcal{Q}_Y\mathcal{Q}_Y^{\mathrm{T}}+\Psi_Y$$

$$\mathcal{Q}_Y\triangleq\mathscr{C}\mathcal{Q}_X\,,\ \Psi_Y=\mathscr{C}\Psi_X\mathscr{C}^{\mathrm{T}}$$

  - 则对于 $\boldsymbol{Y}$，同样含有 k 个因子的因子模型亦成立，其中 $\mathcal{Q}_Y=\mathscr{C}\mathcal{Q}_X$，$\Psi_Y=\mathscr{C}\Psi_X\mathscr{C}^{\mathrm{T}}$.

正交因子模型 (The Orthogonal Factor Model)

- 尺度不变性
  - 在实际应用中，寻找载荷矩阵 $\mathcal{Q}$ 和特殊方差矩阵 Ψ，通常是通过分解 $\boldsymbol{X}$ 的相关矩阵，而非协方差矩阵 Σ 来实现.
  - 这对应于对 $\boldsymbol{X}$ 的线性变换进行因子分析：

$$\boldsymbol{Y} = D^{-1/2}\left(\boldsymbol{X} - \boldsymbol{\mu}\right)$$

  - 目标是尝试找到载荷矩阵 $\mathcal{Q}_Y$ 和特殊方差矩阵 Ψ_Y，使得

$$P = \mathcal{Q}_Y \mathcal{Q}_Y^{\mathrm{T}} + \Psi_Y$$

P 是 $\boldsymbol{X}$ 是的相关矩阵

  - 此时，对因子 $\boldsymbol{F}$ 的解释可以通过以下相关矩阵进行

$$P_{XF} = P_{YF} = \mathcal{Q}_Y$$

正交因子模型 (The Orthogonal Factor Model)

- 尺度不变性

$$\boldsymbol{Y} = D^{-1/2}(\boldsymbol{X} - \boldsymbol{\mu}) \Longrightarrow \mathcal{Q}_Y = D^{-1/2}\,\mathcal{Q}_X$$

$$\Longrightarrow \mathcal{Q}_X = D^{1/2}\,\mathcal{Q}_Y$$

模型的载荷，其中 $\boldsymbol{X}$ 以其原始度量单位表示.

$$\Longrightarrow \Psi_Y = D^{-1/2}\,\Psi_X\,D^{-1/2}$$

$$\Longrightarrow \Psi_X = D^{1/2}\,\Psi_Y\,D^{1/2}$$

模型的特殊方差，其中 $\boldsymbol{X}$ 以其原始度量单位表示.

- 注意：尽管因子分析模型 $\boldsymbol{X} = \mathcal{Q}\boldsymbol{F} + \boldsymbol{U} + \boldsymbol{\mu}$ 具有尺度不变性，但实际估计出的因子可能与尺度有关.

正交因子模型 (The Orthogonal Factor Model)

- 因子载荷不唯一

$$\underset{(p\times 1)}{\boldsymbol{X}} = \underset{(p\times k)}{\mathscr{Q}} \underset{(k\times 1)}{\boldsymbol{F}} + \underset{(p\times 1)}{\boldsymbol{U}} + \underset{(p\times 1)}{\boldsymbol{\mu}}$$

  - 设 $\mathscr{G}$ 为一个正交矩阵: $\mathscr{G}\mathscr{G}^{\mathrm{T}} = \mathscr{G}^{\mathrm{T}}\mathscr{G} = \mathscr{I}_k$.

$$\Longrightarrow \quad \boldsymbol{X} = \mathscr{Q}\boldsymbol{F} + \boldsymbol{U} + \boldsymbol{\mu} = \mathscr{Q}\left(\mathscr{G}\mathscr{G}^{\mathrm{T}}\right)\boldsymbol{F} + \boldsymbol{U} + \boldsymbol{\mu}$$

$$= \left(\mathscr{Q}\mathscr{G}\right)\left(\mathscr{G}^{\mathrm{T}}\boldsymbol{F}\right) + \boldsymbol{U} + \boldsymbol{\mu}$$

  - 这意味着，如果针对 $\boldsymbol{X}$ 的具有因子 $\boldsymbol{F}$ 和载荷 $\mathscr{Q}$ 的 k 个因子的模型成立，那么具有因子 $\mathscr{Q}^{\mathrm{T}}\boldsymbol{F}$ 和载荷 $\mathscr{Q}\mathscr{G}$ 的 k 个因子的模型同样成立.
  - 因子载荷不唯一!
  - 在实际应用中，我们会利用这种不唯一性的优点.

正交因子模型 (The Orthogonal Factor Model)

- 因子载荷不唯一
  - 用正交矩阵左乘向量 $\boldsymbol{F}$ 相当于对坐标轴进行旋转.
  - 新坐标系中第一条轴的方向由正交矩阵的第一行确定.
  - 我们将证明，选择合适的旋转方式，会得到一个更易于解释的载荷矩阵 $\mathscr{Q}\mathscr{G}$.
  - 我们已经看到，因子载荷反映了因子与原始变量之间的相关性.
  - 因此，寻找能够使得因子与不同变量组具有最大相关性的旋转方式是有意义的.

正交因子模型 (The Orthogonal Factor Model)

- 因子载荷不唯一
  - 从数值计算的角度来看，不唯一是一个缺点.
  - 我们必须找到满足以下分解的载荷矩阵 $\mathcal{Q}$ 和特殊方差矩阵 Ψ

$$\Sigma = \mathcal{Q}\mathcal{Q}^{\mathrm{T}} + \Psi$$

  - 由于解的多样性，不存在直接的数值算法能解决这个问题.
  - 一种可行的方法是施加一些选定的约束条件，以便得到分解的唯一解.
  - 然后，我们利用旋转的优点以获得一个更易于解释的解.

正交因子模型 (The Orthogonal Factor Model)

- 因子载荷不唯一
  - 一个显而易见的问题：我们应该施加什么样的约束条件?
  - 通常施加的约束条件是

$$\mathcal{Q}^{\mathrm{T}}\Psi^{-1}\mathcal{Q} \quad \text{是对角阵, 或者} \quad \mathcal{Q}^{\mathrm{T}}D^{-1}\mathcal{Q} \quad \text{是对角阵.}$$

  - 无约束条件时，模型 $\Sigma = \mathcal{Q}\mathcal{Q}^{\mathrm{T}} + \Psi$ 有多少个参数?

$$\mathcal{Q}_{p\times k} \quad \text{有 } p\times k \text{ 个参数}$$

$$\Psi_{p\times p} \quad \text{有 } p \text{ 个参数}$$

  - 因此，我们需要确定 $pk + p$ 个参数的值!
  - 由于要求 $\mathcal{Q}^{\mathrm{T}}\Psi^{-1}\mathcal{Q}$ 或 $\mathcal{Q}^{\mathrm{T}}D^{-1}\mathcal{Q}$ 是对角阵，共有 $\frac{1}{2}k(k-1)$ 个已知条件.

正交因子模型 (The Orthogonal Factor Model)

- 因子载荷不唯一
  - 所以，含有 k 个因子的一个模型的自由度为

$$
\begin{aligned}
d &= (\Sigma \text{ 无约束条件时的参数个数}) - (\Sigma \text{ 有约束条件时的参数个数}) \\
&= \frac{1}{2}p(p+1) - \left(pk + p - \frac{1}{2}k(k-1)\right) \\
&= \frac{1}{2}(p-k)^2 - \frac{1}{2}(p+k)
\end{aligned}
$$

  - 如果 $d < 0$，该模型不确定：$\Sigma = QQ^{\mathrm{T}} + \Psi$ 有无数个解.

    因子模型的参数数量比原始模型的参数数量多.

    或者，因子数量 k 相对于 p “过大”.

正交因子模型 (The Orthogonal Factor Model)

- 因子载荷不唯一
  - 所以，含有 k 个因子的一个模型的自由度为

$$\begin{aligned} d &= (\Sigma \text{ 无约束条件时的参数个数}) - (\Sigma \text{ 有约束条件时的参数个数}) \\ &= \frac{1}{2}p(p+1) - \left(pk + p - \frac{1}{2}k(k-1)\right) \\ &= \frac{1}{2}(p-k)^2 - \frac{1}{2}(p+k) \end{aligned}$$

  - 在某些情况下，$d=0$：该问题存在唯一解 (不考虑旋转情况).

正交因子模型 (The Orthogonal Factor Model)

- 因子载荷不唯一
  - 所以，含有 k 个因子的一个模型的自由度为

$$
\begin{aligned}
d &= (\Sigma\ \text{无约束条件时的参数个数}) - (\Sigma\ \text{有约束条件时的参数个数}) \\
&= \frac{1}{2}p(p+1) - \left(pk + p - \frac{1}{2}k(k-1)\right) \\
&= \frac{1}{2}(p-k)^2 - \frac{1}{2}(p+k)
\end{aligned}
$$

  - 在实际中，我们通常会遇到 $d > 0$ 的情况：方程数量多于参数数量，因此不存在精确解.
  - 此时会使用近似解. 例如，Σ 的一个近似值是 $\mathcal{Q}\mathcal{Q}^{\mathrm{T}} + \Psi$.
  - 参数估计的方法将在下一节介绍.

正交因子模型 (The Orthogonal Factor Model)

- 因子载荷不唯一
  - 评估自由度 d 尤为重要，因为它能让我们初步了解在因子模型中有望识别出的因子数量上限.

$$d = \frac{1}{2}(p-k)^2 - \frac{1}{2}(p+k) \xrightarrow{p=4,\, k=1} d = 2 > 0$$

$\Longrightarrow$ 只有单一因子模型能给出一个近似解.

$$d = \frac{1}{2}(p-k)^2 - \frac{1}{2}(p+k) \xrightarrow{p=4,\, k=2} d = -1 < 0$$

$\Longrightarrow$ 我们无法确定一个包含两个因子的因子模型.

正交因子模型 (The Orthogonal Factor Model)

- 因子载荷不唯一
  - 评估自由度 d 尤为重要，因为它能让我们初步了解在因子模型中有望识别出的因子数量上限.

$$d = \frac{1}{2}(p-k)^2 - \frac{1}{2}(p+k) \xrightarrow{p=6,\ k=1} d = 6 > 0$$

$$\xrightarrow{p=6,\ k=2} d = 4 > 0$$

$\Longrightarrow$ 具有 1 个和 2 个因子的模型可提供近似解.

$$d = \frac{1}{2}(p-k)^2 - \frac{1}{2}(p+k) \xrightarrow{p=6,\ k=3} d = 0$$

$\Longrightarrow$ 具有三个因子的模型会有唯一解 (不考虑旋转的情况).

$$d = \frac{1}{2}(p-k)^2 - \frac{1}{2}(p+k) \xrightarrow{p=6,\ k=4} d = -3 < 0$$

$\Longrightarrow$ 不允许使用具有 4 个或更多因子的模型.

正交因子模型 (The Orthogonal Factor Model)

- 例: 设 $p=3$ 及 $k=1$, 则

$$d=\frac{1}{2}(p-k)^2-\frac{1}{2}(p+k)=0$$

$$\Sigma=\mathcal{Q}\mathcal{Q}^{\mathrm{T}}+\Psi \implies \Sigma=\begin{pmatrix}\sigma_{11}&\sigma_{12}&\sigma_{13}\\ \sigma_{21}&\sigma_{22}&\sigma_{23}\\ \sigma_{31}&\sigma_{32}&\sigma_{33}\end{pmatrix}=\begin{pmatrix}q_1\\ q_2\\ q_3\end{pmatrix}\begin{pmatrix}q_1&q_2&q_3\end{pmatrix}+\begin{pmatrix}\psi_{11}&&\\ &\psi_{22}&\\ &&\psi_{33}\end{pmatrix}$$

$$\implies \Sigma=\begin{pmatrix}\sigma_{11}&\sigma_{12}&\sigma_{13}\\ \sigma_{21}&\sigma_{22}&\sigma_{23}\\ \sigma_{31}&\sigma_{32}&\sigma_{33}\end{pmatrix}=\begin{pmatrix}q_1^2+\psi_{11}&q_1q_2&q_1q_3\\ q_1q_2&q_2^2+\psi_{22}&q_2q_3\\ q_1q_3&q_2q_3&q_3^2+\psi_{33}\end{pmatrix}$$

$$\implies q_1^2=\frac{\sigma_{12}\sigma_{13}}{\sigma_{23}}\ ;\quad q_2^2=\frac{\sigma_{12}\sigma_{23}}{\sigma_{13}}\ ;\quad q_3^2=\frac{\sigma_{13}\sigma_{23}}{\sigma_{12}}$$

$$\psi_{11}=\sigma_{11}-q_1^2\ ;\quad \psi_{22}=\sigma_{22}-q_2^2\ ;\quad \psi_{33}=\sigma_{33}-q_3^2$$

- 在这种特定情况 $(k=1)$ 下，唯一的旋转由 $\mathcal{G}=-1$ 定义，所以，载荷的其它解方式由 $-\mathcal{Q}$ 给出.

正交因子模型 (The Orthogonal Factor Model)

- 例: 设 $p=3$ 及 $k=1$, 则

$$d=\frac{1}{2}(p-k)^2-\frac{1}{2}(p+k)=0$$

$$\Sigma = QQ^{\mathrm{T}}+\Psi \implies \Sigma=\begin{pmatrix}\sigma_{11} & \sigma_{12} & \sigma_{13}\\ \sigma_{21} & \sigma_{22} & \sigma_{23}\\ \sigma_{31} & \sigma_{32} & \sigma_{33}\end{pmatrix}=\begin{pmatrix}q_1\\ q_2\\ q_3\end{pmatrix}\begin{pmatrix}q_1 & q_2 & q_3\end{pmatrix}+\begin{pmatrix}\psi_{11} & & \\ & \psi_{22} & \\ & & \psi_{33}\end{pmatrix}$$

$$\implies \Sigma=\begin{pmatrix}\sigma_{11} & \sigma_{12} & \sigma_{13}\\ \sigma_{21} & \sigma_{22} & \sigma_{23}\\ \sigma_{31} & \sigma_{32} & \sigma_{33}\end{pmatrix}=\begin{pmatrix}q_1^2+\psi_{11} & q_1q_2 & q_1q_3\\ q_1q_2 & q_2^2+\psi_{22} & q_2q_3\\ q_1q_3 & q_2q_3 & q_3^2+\psi_{33}\end{pmatrix}$$

$$\implies q_1^2=\frac{\sigma_{12}\sigma_{13}}{\sigma_{23}}\ ;\quad q_2^2=\frac{\sigma_{12}\sigma_{23}}{\sigma_{13}}\ ;\quad q_3^2=\frac{\sigma_{13}\sigma_{23}}{\sigma_{12}}$$

$$\psi_{11}=\sigma_{11}-q_1^2\ ;\quad \psi_{22}=\sigma_{22}-q_2^2\ ;\quad \psi_{33}=\sigma_{33}-q_3^2$$

- 该解决方案可能是唯一的 (旋转情况除外)，但从统计学上无法解释这一角度而言，它并不恰当.

正交因子模型 (The Orthogonal Factor Model)

- 例: 现在假设 $p=2$ 及 $k=1$, 则

$$d=\frac{1}{2}(p-k)^2-\frac{1}{2}(p+k)=-1<0$$

$$\Sigma = \mathcal{Q}\mathcal{Q}^{\mathrm{T}}+\Psi \quad\Longrightarrow\quad \Sigma=\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix}=\begin{pmatrix}q_1\\ q_2\end{pmatrix}\begin{pmatrix}q_1 & q_2\end{pmatrix}+\begin{pmatrix}\psi_{11} & \\ & \psi_{22}\end{pmatrix}=\begin{pmatrix}q_1^2+\psi_{11} & q_1q_2\\ q_1q_2 & q_2^2+\psi_{22}\end{pmatrix}$$

$$\Longrightarrow\quad q_1=\alpha\,;\quad q_2=\frac{\rho}{\alpha}\,;\quad \psi_{11}=1-\alpha^2\,;\quad \psi_{22}=1-\left(\frac{\rho}{\alpha}\right)^2;\ \forall\alpha\in(\rho\,,\,1)$$

有无数个解!

即便在存在唯一解 $(d=0)$ 的情况下，该解也可能与统计学的解释相悖.