

主要研究内容

- 一、管理研究概论
- 二、研究设计
- 三、数据收集
- 四、数据处理方法
- 五、研究结果呈现

管理研究方法

—— 探索性因子分析

-
- 因子分析的概念
 - 因子分析的步骤
 - ◆ 设计因子分析
 - ◆ 因子分析的隐含假设
 - ◆ 获取因子并评估拟合度
 - ◆ 解释因子
 - ◆ 因子分析的验证
 - ◆ 因子分析结果的其他用处

学习目标

学完这一部分内容，你要能够：

1. 了解因子分析的基本步骤；
 2. 主成份法和因子分析的区别；
 3. 确定进行因子分析所需的样本大小；
 4. 选择因子旋转方法；
 5. 确定抽取因子的个数；
 6. 使用因子分析的结果。
-

因子分析的概念

因子分析是一种相关分析技术（interdependent technique），其主要目的是界定分析所涉及变量的隐含结构。

因子分析的概念

V1	.611	.250
V2	.614614	-.446- 446
V3	.750	.312
V4	.540	.112
V5	.250	.550
V6	-.321	.659
V7	.222	.450
V8	.118	.780

Data reduction



Data reduction



因子得分:

$$\text{Factor 1} = .611*V1+.614*V2+.750*V3+...+.118*V8$$

$$\text{Factor 2} = .250*V1-.446*V2+.312*V3+...+.780*V8$$

因子分析的步骤

1. 设计因子分析
2. 因子分析的隐含假设
3. 获取因子并评估拟合度
4. 解释因子
5. 因子分析的验证
6. 因子分析结果的其他用处

1.设计因子分析

- 什么样的变量可以进行因子分析？
 - 至少要求能够计算所有变量的相关系数。
- 应该包含多少个变量？
 - 每个目标因子至少应具有5个变量与其对应。

样本量

- 样本量应至少达到50。样本量和变量数的比率应大于10:1
- 但是，样本量和变量数的比率要求不是绝对的，在下面三种情况下，这一要求可降低：
 - 题项共同度（community）高（达到0.8或更高）。
 - 没有交叉载荷（没有在两个以上的因子上载荷高于0.32的情况）。
 - 一个因子包含的变量至少达到三个。

2. 因子分析的隐含假设

- 在进行因子分析前，必须确保潜在的因子结构具有坚实的理论支持。

实例： *Chen, Y., Friedman, R., Yu, E., Fang, W., & Lu, X. 2009. Supervisor-subordinate guanxi: Developing a three-dimensional model and scale. Management and Organization Review, 5(3): 375-399.*

we selected six to 10 statements from each category, choosing the most frequently mentioned items. ...A principal axis factor analysis with promax rotation resulted in three factors with eigenvalues greater than one, ...We deleted items with cross-loadings on factors other than the intended factor...**we also dropped the items with content that was inconsistent with other items loading on the same factor.**

统计假设

- Bartlett test of sphericity
 - 若显著则表明变量间存在显著的相关性。
- Measure of sampling adequacy (KMO)
 - 采用Kaiser-Meyer-Olkin (KMO) 测量。要求总体KMO和单个变量的KMO均达到0.50。若没有达到这一要求，则去除KMO最低的变量后重新计算总体和单个变量的KMO。如此反复，直至达到KMO要求。

3. 获取因子并评估拟合度

- 主成份分析 (principal component method) *vs.* 因子分析 (factor analysis)
 - 主成份方法不考虑潜变量引起的隐含结构。当因子之间不相关，共同度 (communities) 中等时，主成份分析会虚增被解释的方差。

1	2.6232	2121.8558	2121.8558	1	2.0702	1717.2542	1717.2542
	623	55	55		070	54	54
2	2.323	19.356	41.211	2	1.956	16.301	33.555
3	1.9121	1515.9309	5757.1411	3	1.4111	411111.7607	4545.3163
	912	30	41		60	16	

虚高的解释方差

数据来源：在宁农民工城市融合问题调研，上司-下属关系量表的调研结果

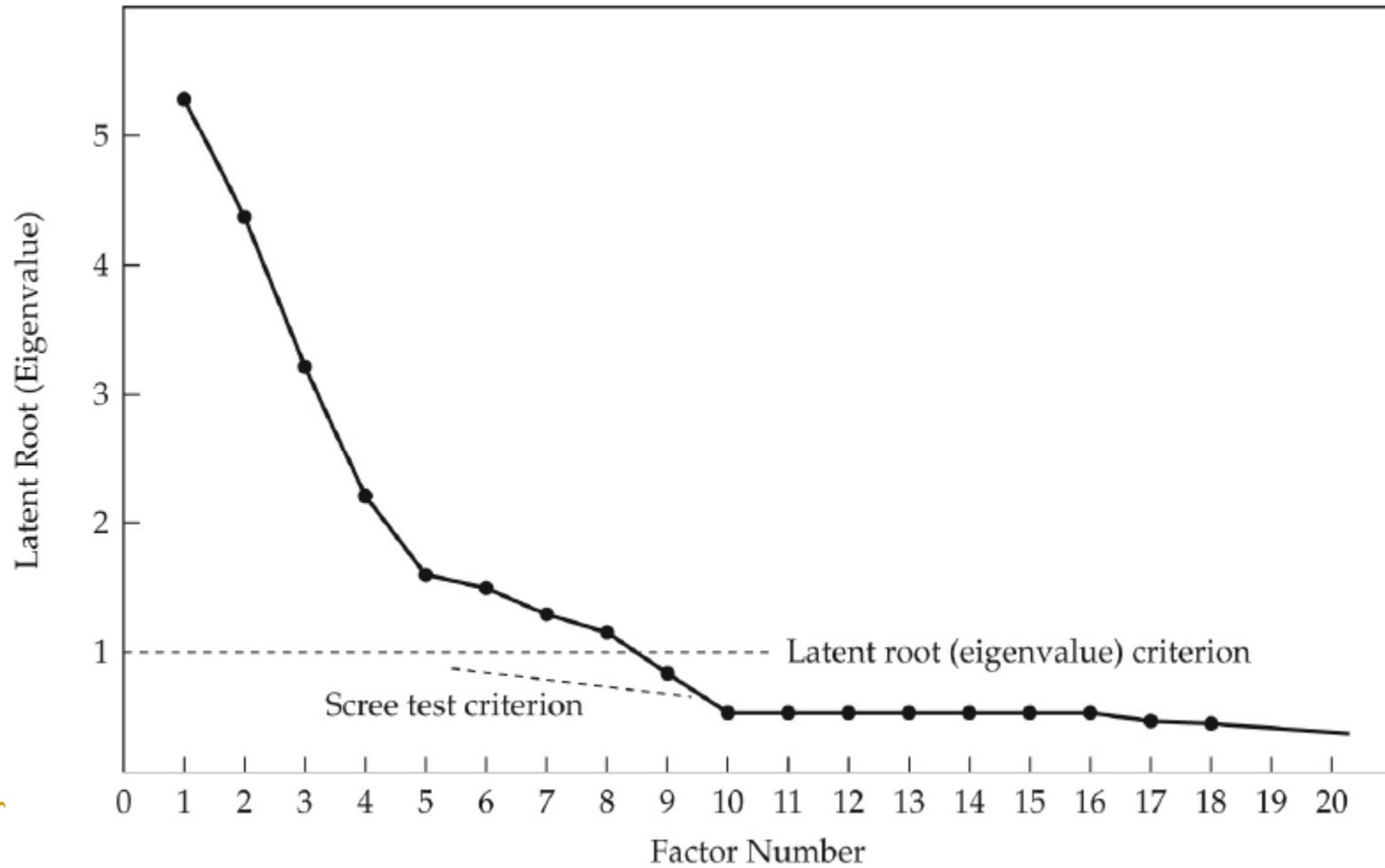
因子分析方法

- 如果数据基本符合正态分布，那么最大似然法（maximum likelihood）是最佳选择，因为这种方法采用了多种拟合优度指数来评估分析结果，并且允许检验因子之间的相关系数和因子载荷的显著性。如果正态分布假设被严重违反，那么主成份因子方法（principal factor methods，在SPSS中称为principal axis factors）是不错的选择（Fabrigar, Wegener, MacCallum, and Strahan, 1999）。

抽取的因子个数

- 掌握因子结构的知识
 - 实例：建构的理论结构
- 累计方差解释率达到一定比率，通常为50-60%
- 特征根（eigenvalue） > 1 或寻找碎石图（scree plot）曲线的折断点

碎石图

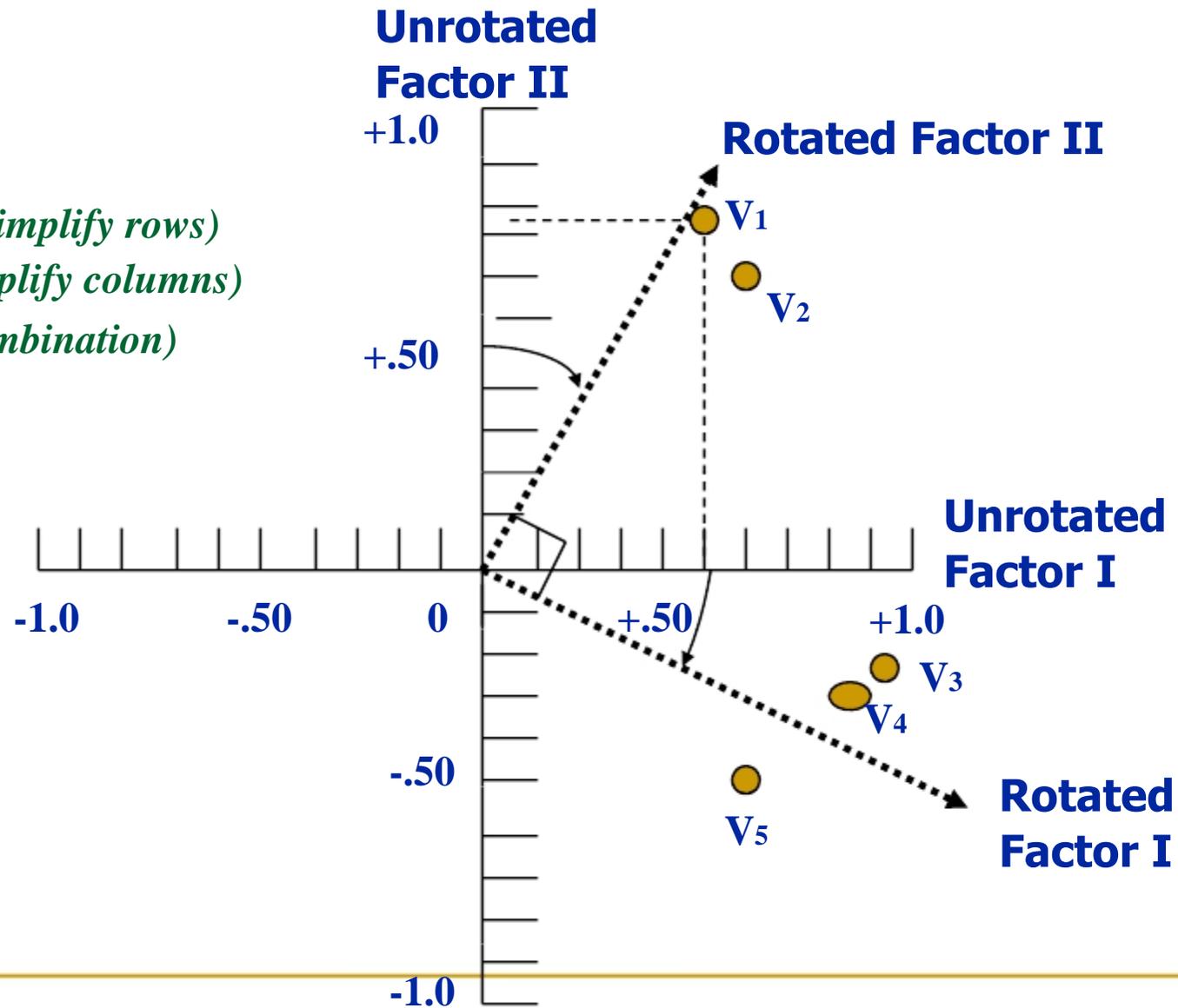


因子旋转 (factor rotation)

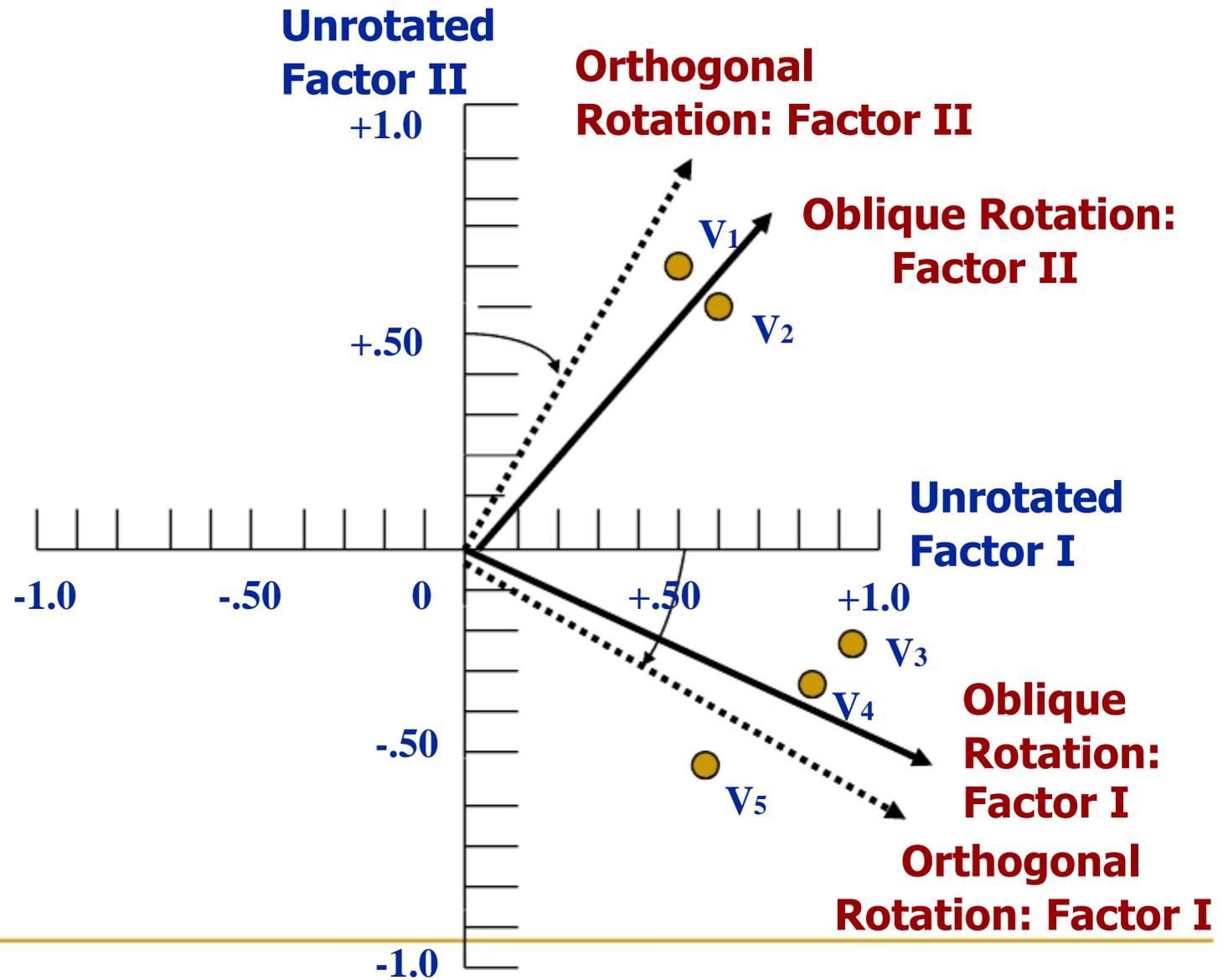
因子旋转是指使坐标轴围绕原点进行旋转，使坐标轴达到一个新的位置，其目的是改善因子拟合数据的情况，使拟合得到的因子更简洁，并且在理论上更具有意义。

正交旋转 (Orthogonal Factor Rotation)

- *Quartimax (simplify rows)*
- *Varimax (simplify columns)*
- *Equimax (combination)*



斜交旋转 (Oblique Factor Rotation)



选择因子旋转方法

■ 正交旋转

- 当因子分析的主要目的是简化数据（data reduction）时，正交旋转是适合采用的。

■ 斜交旋转

- 当我们需要识别构念的结构时，斜交旋转比正交旋转更合适，因为在现实世界中，一个构念的几个维度完全不相关的情况几乎是不存在的。

评估因子载荷 (factor loadings) 的统计显著性

Factor Loadingg

**Sample Size Needed
for Significance***

.30	350
.35	250
.40	200
.45	150
.50	120
.55	100
.60	85
.65	70
.70	60
.75	50

***Significance is based on a .05 significance level (α), a power level of 80 percent, and standard errors assumed to be twice those of conventional correlation coefficients.**

评估因子载荷的现实显著性 (practical significance)

- 因子载荷如果介于 ± 0.30 和 ± 0.40 ，那么可以认为达到了解释数据背后结构的基本水平。
- 如果因子载荷高于 ± 0.50 ，可以认为其具备了现实显著性。
- 因子载荷高于 ± 0.70 被认为相当理想，是因子分析追求的目标。

4. 解释因子

	<i>Personal-life inclusion</i>		<i>Deference to supervisor</i>	
V1: 我的上司会请我帮他处理一些私事	.764	.242	.643	>.50
V2: 我和上司在非工作时间会有一些和工作无关的社交活动	.754	.123	.583	>.40
V3: 我熟悉上司的家人并且和他们有些个人联系	.817	.077	.673	<.40
V4: 为了实现上司的目标, 我愿意放弃自己的目标	.270	.686	.543	
V5: 即使和上司观点不一致, 我也会支持他	-.029	.813	.661	
V6: 为了上司的利益, 我愿意放弃自己的利益	.219	.733	.585	

Note: Extraction Method: Principal Component Analysis.
 Rotation Method: Varimax with Kaiser Normalization.

未见交叉载荷
 (*cross-loading*)

因子模型的重新设定

Chen, Y., Friedman, R., Yu, E., Fang, W., & Lu, X. 2009. Supervisor-subordinate guanxi: Developing a three-dimensional model and scale. Management and Organization Review, 5(3): 375-399.

Table 1. Exploratory factor analysis results of 24 items

Retained items			
1. My supervisor and I always share thoughts, opinions, and feelings toward work and life.	0.46	0.26	0.14
2.2 I feel easy and comfortable when I communicate with my supervisor.	0.750 75	0.110 11	0.030 03
.....			
Excluded items			
13. I have a good personal relationship with my supervisor.	0.56	0.49	-0.15
17. I greatly respect and defer to my supervisor	0.88	0.20	0.00
.....			

not load on the factor for which it was originally intended the item did

5. 因子分析的验证

- 进行确认性因子分析（confirmatory factor analysis）
- 评估因子结构的稳定性
 - 样本规模允许情况下，分半进行检测
- 识别异常值（influential observations）
 - 保留或去除异常值，评估因子分析结果是否一致

6. 因子分析结果的其他用处

V1:我的上司会请我帮他处理一些私事	.764764	.242242	.643643
V2:我和上司在非工作时间会有一些和工作无关的社交活动	.754	.123	.583
V3V3:我熟悉上司的家人并且和他们有些个人联系	.817817	.077077	.673673
V4:为了实现上司的目标, 我愿意放弃自己的目标	.270	.686	.543
V5:即使和上司观点不一致, 我也会支持他	-.029.029	.813	.661
V6:为了上司的利益, 我愿意放弃自己的利益	.219	.733	.585

The single surrogate variable:

$$\text{Factor 1} = V3$$

Summated Scale:

$$\text{Factor 1} = \sum \text{factor loading}_i * V_i \quad \text{or} \quad (\sum V_i) / 3 \quad (i = 1, 2, 3)$$

Factor Score:

$$\text{Factor 1} = \sum \text{factor loading}_i * V_i \quad (i = 1, 2, 3, 4, 5, 6)$$

阅读文献

- Hair Jr., J. F., Black, W. C., Babin, B. J., and Anderson, R. E. 2010. *Multivariate Data Analysis (7th edition)*. Pearson Prentice Hall. Chapter 3.
- Costello, A. B. & Osborne, J. W. 2005. Best practice in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment, Research & Evaluation*, 10(7).
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. 1999. Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3): 272-299.