Factor Analysis

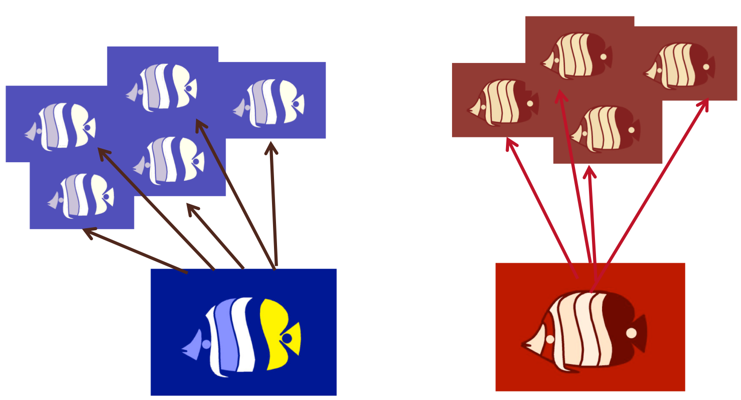
CLPS 2908
Lecture 12

3/12/2019

Basic Idea of Factor Analysis

Observed variables are taken to be **indicators** of underlying, unobservable **latent constructs**.

Constructs can be estimated from swarms of observed variables, which are **caused** by the constructs.



Model Assumptions

- Explicit error model on variables
  From classical test theory: $X = T + e$
- Common factors model: Every variable is the result of loadings on common factors and error:

$$\underset{p\times 1}{y_j} = \underset{p\times r}{l_{jk}} \cdot \underset{r\times 1}{f_k} + \underset{p\times 1}{\varepsilon_j}$$

y = vector of j mean-deviated variables for a given individual

$$\underset{p\times n}{\mathbf{Y}} = \underset{p\times r}{\mathbf{L}} \cdot \underset{r\times n}{\mathbf{F}} + \underset{p\times n}{\varepsilon}$$

$\mathbf{Y}$ = matrix of j mean-deviated variables for entire sample

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

No longer a mere mathematical transformation as in PCA, where: **C = X V** **R = V Λ V'**

Ψ = var/cov matrix of errors ε_i for each variable

Factor Analysis Procedure

- GOAL: Extract and remove error
  $s_x^2 - e^2 \Rightarrow s_T^2$, from $X - e \Rightarrow T$
- PCA on a "reduced" correlation matrix (minus **E)**
- Before PCA, estimate each variable's error variance as *variance not accounted for by other variables in the set* (➛ multiple regressions for each variable: R_j^2)
  - Assume that observed variables are sampled from universe of variables of a given construct
- Starting communality h^2 = common variance $(1-e^2)$
- Starting matrix: $\mathbf{R_{xx}} - \mathbf{E} = \mathbf{R}_{\text{reduced}}$

Based on multiple regression of each variable on all others

$$\begin{bmatrix} s_1^2 & & & \\ r_{21} & s_2^2 & & \\ r_{31} & r_{32} & s_3^2 & \\ r_{41} & r_{42} & r_{43} & s_4^2 \end{bmatrix} - \begin{bmatrix} e_1^2 & 0 & 0 & 0 \\ 0 & e_2^2 & 0 & 0 \\ 0 & 0 & e_3^2 & 0 \\ 0 & 0 & 0 & e_4^2 \end{bmatrix} = \begin{bmatrix} h_1^2 & & & \\ r_{21} & h_2^2 & & \\ r_{31} & r_{32} & h_3^2 & \\ r_{41} & r_{42} & r_{43} & h_4^2 \end{bmatrix}$$

$$\mathbf{R} \quad - \quad \mathbf{E} \quad = \quad \mathbf{R}_{\text{reduced}}$$

Correlations (off-diagonals) are unchanged because they do not contain error variances (defined as unique variance that each variable does not share with other variables)

Factor Analysis Process

1 Create reduced **R** (from multiple regressions: R_j^2)

2 Extract, as usual, PCs (= "factors") from $\mathbf{R}_{\text{reduced}}$

3 On the basis of r initial factors,

re-estimate communalities, $\hat{h}_j^2$

and all variable intercorrelations, $\hat{r}_{jj}$,

($\rightarrow$ $\mathbf{R}_{\text{reproduced}}$)

$$\begin{bmatrix} \hat{h}_1^2 & & & \\ \hat{r}_{21} & \hat{h}_2^2 & & \\ \hat{r}_{31} & \hat{r}_{32} & \hat{h}_3^2 & \\ \hat{r}_{41} & \hat{r}_{42} & \hat{r}_{43} & \hat{h}_4^2 \end{bmatrix}$$

4 Now extract factors again from $\mathbf{R}_{\text{reproduced}}$.

5 Repeat extraction and re-estimation iteratively until two succeeding $\mathbf{R}_{\text{reproduced}}$ do not differ

FA Syntax

```
FACTOR VAR = X1 to X25
 /plot = eigen rotation
 /format = sort blank (.30)
 /print = default  fscore
 /criteria = factors(r)
 /extraction = paf
 /rotation = varimax [oblimin].
```

Oblique Rotation

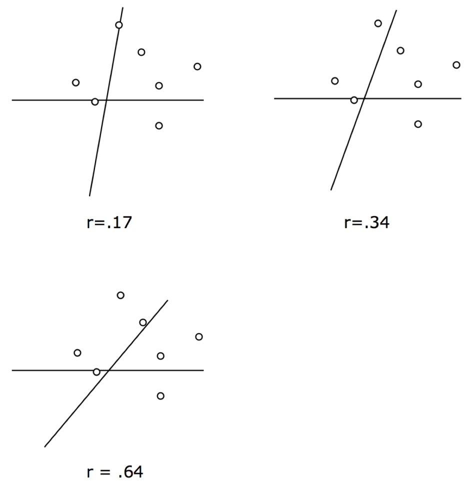
- Theoretically, latent constructs may be correlated.
- If FA tries to capture correlated latent constructs, correlated factors are appropriate.

PCA: $\mathbf{R}_{xx} = \mathbf{A}\,\mathbf{A}'$

FA: $\mathbf{R}_{xx} = \mathbf{L}\,\Phi\,\mathbf{L}' + \Psi$

Factor intercorrelation matrix

Oblique Rotation: Axes Only



PCA

unrotated

λ_1	6.751
λ_2	4.552
λ_3	3.210
λ_4	2.406
λ_5	1.906

Communalities

	Initial	Extraction
DISTA	1.000	.533
TALKA	1.000	.684
CAREL	1.000	.509
HARDW	1.000	.524
ANXIO	1.000	.661
AGREE	1.000	.550
TENSE	1.000	.722
KIND	1.000	.589
OPPOS	1.000	.536
RELAX	1.000	.680
DISOR	1.000	.589
OUTGO	1.000	.752
APPRO	1.000	.456
SHY	1.000	.631
DISCI	1.000	.497
HARSH	1.000	.580
PERSE	1.000	.471
FRIEN	1.000	.634
WORRY	1.000	.673
RESPO	1.000	.606
CONRY	1.000	.649
SOCIA	1.000	.640
LAZY	1.000	.553
COOPE	1.000	.551
QUIET	1.000	.730
ORGAN	1.000	.604
CRITI	1.000	.546
LAX	1.000	.422
LAIDB	1.000	.588
WITHD	1.000	.695
GIVUP	1.000	.458
EASYG	1.000	.511

Extraction Method: Principal Component Analysis.

Communalities

	Initial	Extraction
DISTA	.512	.450
TALKA	.620	.622
CAREL	.554	.438
HARDW	.531	.459
ANXIO	.578	.573
AGREE	.509	.432
TENSE	.648	.679
KIND	.555	.511
OPPOS	.445	.409
RELAX	.588	.625
DISOR	.723	.534
OUTGO	.733	.729
APPRO	.425	.364
SHY	.628	.566
DISCI	.510	.414
HARSH	.472	.474
PERSE	.467	.394
FRIEN	.613	.578
WORRY	.622	.595
RESPO	.560	.553
CONRY	.530	.577
SOCIA	.617	.579
LAZY	.532	.494
COOPE	.503	.467
QUIET	.671	.692
ORGAN	.716	.556
CRITI	.397	.428
LAX	.390	.324
LAIDB	.553	.480
WITHD	.646	.653
GIVUP	.501	.400
EASYG	.498	.413

Extraction Method: Principal Axis Factoring.

FA

unrotated

λ_1	6.297
λ_2	4.049
λ_3	2.781
λ_4	1.938
λ_5	1.397

Orthogonal Rotation: PCA vs. FA

PCA

Rotated Component Matrix[a]

	Component 1	2	3	4	5
ORGAN	-.773				
DISOR	.750				
RESPO	-.733			.244	
LAZY	.702				
DISCI	-.679				
HARDW	-.654				
CAREL	.638				.301
PERSE	-.587			.239	
GIVUP	.495	.368	.232		
LAX	.488		-.272	.295	
OUTGO		-.836		.211	
QUIET		.812			
TALKA		-.791			
WITHD		.764			.264
SOCIA		-.751		.257	
SHY		.738			
DISTA		.646			.316
TENSE			.793		.263
WORRY			.789		
ANXIO			.736		.221
RELAX			-.728	.356	
LAIDB	.280		-.645	.285	
EASYG			-.503	.472	
AGREE				.722	
KIND	-.306			.675	
COOPE				.655	-.258
APPRO		-.233		.603	
FRIEN		-.502		.583	
CONRY					.768
HARSH				-.200	.726
OPPOS					.712
CRITI	-.243				.666

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

FA

Rotated Factor Matrix[a]

	Factor 1	2	3	4	5
OUTGO	-.821			.214	
QUIET	.790				
TALKA	-.756				
WITHD	.734				.261
SOCIA	-.710			.255	
SHY	.699				
DISTA	.589				.283
ORGAN		-.740			
DISOR		.715			
RESPO		-.703		.228	
LAZY		.661			
DISCI		-.625			
HARDW		-.617			
CAREL		.595			.271
PERSE		-.547		.202	
GIVUP	.345	.458	.210		
LAX		.451	-.225	.244	
TENSE			.767		.262
WORRY			.741		
ANXIO			.687		.215
RELAX			-.683	.360	
LAIDB		.282	-.567	.271	
EASYG			-.437	.428	
AGREE				.636	
KIND		-.299		.620	
COOPE				.591	-.248
FRIEN	-.479			.553	
APPRO	-.231			.523	
CONRY					.718
HARSH					.648
OPPOS					.616
CRITI		-.219			.578

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.

Oblique vs. Orthogonal Rotation

FA

oblique

$\lambda_1 = 4.97$
$\lambda_2 = 4.56$
$\lambda_3 = 2.90$
$\lambda_4 = 2.79$
$\lambda_5 = 3.53$

Pattern Matrix[a]

	Factor 1	2	3	4	5
OUTGO	-.831				
QUIET	.805			.260	
TALKA	-.789				
WITHD	.727		.232		
SOCIA	-.718				
SHY	.696				
DISTA	.587		.268		
ORGAN		.752			
RESPO		.724		.218	
DISOR		-.718			
LAZY		-.653			
DISCI		.651			
HARDW		.610			
CAREL		-.586	.244		
PERSE		.542			
GIVUP	.287	-.441			
LAX		-.416		.254	.228
CONRY			.720		
HARSH			.656		
OPPOS			.625		
CRITI		.275	.587		
AGREE				.629	
KIND		.286		.599	
COOPE			-.214	.555	
FRIEN	-.454			.505	
APPRO				.487	
TENSE					-.758
WORRY					-.752
RELAX				.306	.696
ANXIO				.231	-.685
LAIDB		-.213		.241	.587
EASYG				.389	.432

Extraction Method: Principal Axis Factoring.
Rotation Method: Oblimin with Kaiser Normalization.

FA

orthog.

$\lambda_1 = 4.36$
$\lambda_2 = 4.30$
$\lambda_3 = 2.95$
$\lambda_4 = 2.57$
$\lambda_5 = 2.29$

Rotated Factor Matrix[a]

	Factor 1	2	3	4	5
OUTGO	-.821			.214	
QUIET	.790				
TALKA	-.756				
WITHD	.734				.261
SOCIA	-.710			.255	
SHY	.699				
DISTA	.589				.283
ORGAN		-.740			
DISOR		.715			
RESPO		-.703		.228	
LAZY		.661			
DISCI		-.625			
HARDW		-.617			
CAREL		.595			.271
PERSE		-.547		.202	
GIVUP	.345	.458	.210		
LAX		.451	-.225	.244	
TENSE			.767		.262
WORRY			.741		
ANXIO			.687		.215
RELAX			-.683	.360	
LAIDB		.282	-.567	.271	
EASYG			-.437	.428	
AGREE				.636	
KIND		-.299		.620	
COOPE				.591	-.248
FRIEN	-.479			.553	
APPRO	-.231			.523	
CONRY					.718
HARSH					.648
OPPOS					.616
CRITI		-.219			.578

Extraction Method: Principal Axis Factoring.
Rotation Method: Varimax with Kaiser Normalization.

A Few Guidelines

- Cases per variable: 5 minimum, 10 better
- Sample size total: > 60 if communalities are large (h^2 > .60); > 200 minimum if communalities are smaller (h^2 < .50)
- Variables per factor: > 4 (more if more factors)
- Typical extraction: $\lambda_i > 1$ (for PCA)
  - Oddity that initial r ($< p$) is derived from $\mathbf{R}_{\text{original}}$
  - FA reduces variance ➟ λs lower ➟ extract fewer factors
  - Parallel analysis recommended

Percentages of Convergent and Admissible Solutions in Our Monte Carlo Study

	Ratio of variables to factors and communality level	Sample size 60	100	200	400
	10:3 ratio				
.2 to .4	Low communality	74.6	78.7	95.2	99.0
.2 to .8	Wide communality	99.0	98.0	99.0	98.0
.6 to .8	High communality	100	100	100	100
	20:3 ratio				
	Low communality	87.0	97.1	100	100
	Wide communality	100	100	100	100
	High communality	100	100	100	100
	20:7 ratio				
	Low communality	4.1	15.8	45.9	80.7
	Wide communality	16.5	42.9	72.5	81.3
	High communality	39.7	74.6	91.7	97.1

PCA vs. FA

PCA	FA
Mathematical transformation of scores $\underset{n \times p}{\mathbf{PC}} = \underset{n \times p}{\mathbf{X}} \; \underset{p \times p}{\mathbf{V}}$	Reinterpration of variables $\underset{p \times n}{\mathbf{Y}} = \underset{p \times r}{\mathbf{L}} \cdot \underset{r \times n}{\mathbf{F}} + \underset{p \times n}{\varepsilon}$
Data reduction via composites	Find underlying latent structure
Spectral decomposition of original **R**	Spectral decomposition of *reduced* **R** (h^2 estimates)
Redistribute variance to find fewest composites	Iteratively reestimates communalities to infer latent factors
Can reconstruct 100% variance	Never reconstructs 100% variance
No explicit error model	Explicit error model ($x = T + e$)
Reversible: $\mathbf{R} = \mathbf{V}\boldsymbol{\Lambda}\mathbf{V}'$; $\boldsymbol{\Lambda} = \mathbf{V}'\mathbf{R}\,\mathbf{V}$	Not reversible ($\mathbf{R}_{orig}$ is abandoned)
PC scores are computed	Factor scores are estimates
Rotation should be orthogonal	Rotation is flexible $\mathbf{R}_{xx} = \mathbf{L}\,\boldsymbol{\Phi}\,\mathbf{L}' + \boldsymbol{\Psi}$