Preliminaries
○○○○○○○○
The basic concepts
○○
Principal Components
○○○○
○○○
Rotations and Transformations
References

Psychology 371: Personality Research
Psychometric Theory – Data reduction

William Revelle

Department of Psychology
Northwestern University
Evanston, Illinois USA

NORTHWESTERN
UNIVERSITY

May, 2016

Preliminaries
00000000

The basic concepts
oo

Principal Components
oooo
ooo

Rotations and Transformations

References

Outline

Models of data

(MacCallum, 2004) "A factor analysis model is not an exact representation of real-world phenomena.
Always wrong to some degree, even in population.
At best, model is an approximation of real world."
Box (1979): "Models, of course, are never true, but fortunately it is only necessary that they be useful. For this it is usually needful only that they not be grossly wrong."
Tukey (1961): "In a single sentence, the moral is: Admit that complexity always increases, first from the model you fit to the data, thence to the model you use to think and plan about the experiment and its analysis, and thence to the true situation."
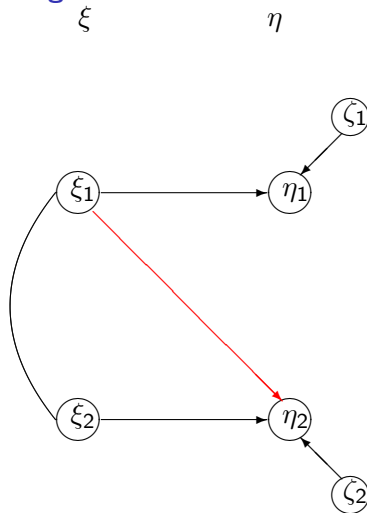
(From MacCallum, 2004); http://www.fa100.info/maccallum2.pdf

Preliminaries
○●○○○○○○

The basic concepts
○○

Principal Components
○○○○
○○○

Rotations and Transformations

References

Observed Variables

X Y

X_1 Y_1

X_2 Y_2

X_3 Y_3

X_4 Y_4

X_5 Y_5

X_6 Y_6

Latent Variables

ξ $\qquad\qquad\qquad\qquad$ η

ξ_1 $\qquad\qquad\qquad\qquad$ η_1

ξ_2 $\qquad\qquad\qquad\qquad$ η_2

Theory: A regression model of latent variables

A measurement model for X

δ $\hspace{4em}$ X $\hspace{4em}$ ξ

Preliminaries
○○○○○●○○

The basic concepts
○○

Principal Components
○○○○
○○○

Rotations and Transformations

References
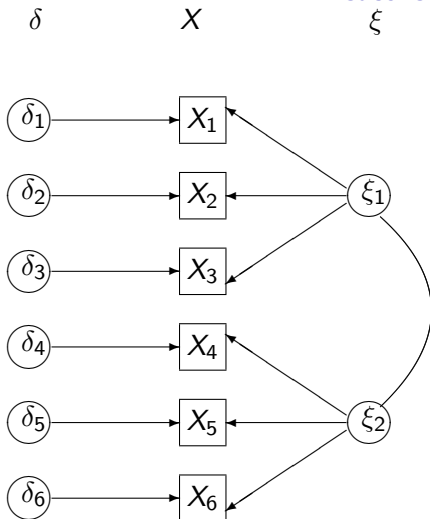
A measurement model for Y
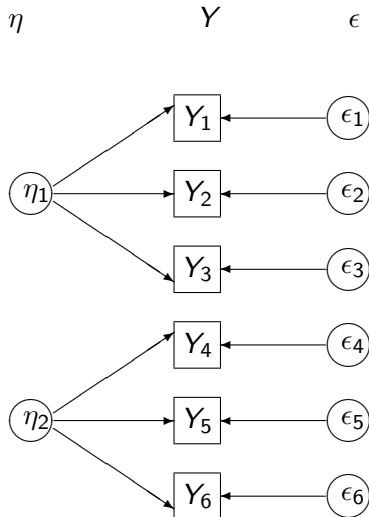
A complete structural model

δ X ξ η Y ϵ
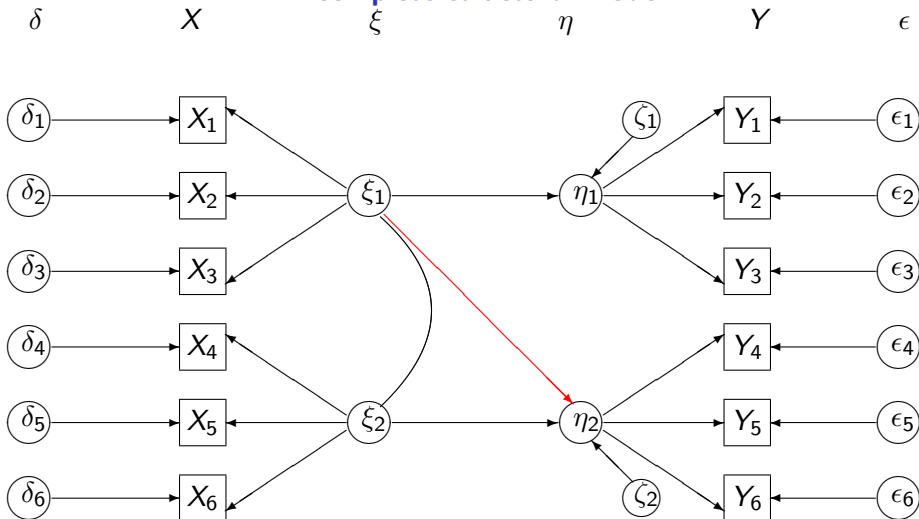
Various measurement models

1. Observed variables models
    - Singular Value Decompostion
    - Eigen Value – Eigen Vector decomposition
    - Principal Components
    - First k principal components as an approximation
2. Latent variable models
    - Factor analysis
3. Interpretation of models
    - Choosing the appropriate number of components/factors
    - Transforming/rotating towards interpretable structures

Factor Analysis/Components Analysis/Cluster Analysis

1. Data simplification and Ockham's Razor: "do not multiple entities beyond necessity"
2. Can we describe a data set with a simpler representation of the data.
3. Is it possible to combine subjects and or variables that are redundant?
4. Or almost redundant (without losing very much information)
5. This is a problem in projective geometry. Can we project from a high dimensional space into a lower order space.

Preliminaries
00000000

The basic concepts
●○

Principal Components
0000
000

Rotations and Transformations

References

An example correlation matrix

Consider the following correlation matrix

```
    V1   V2   V3   V4   V5   V6
V1 1.00 0.72 0.63 0.54 0.45 0.36
V2 0.72 1.00 0.56 0.48 0.40 0.32
V3 0.63 0.56 1.00 0.42 0.35 0.28
V4 0.54 0.48 0.42 1.00 0.30 0.24
V5 0.45 0.40 0.35 0.30 1.00 0.20
V6 0.36 0.32 0.28 0.24 0.20 1.00
```

Is it possible to model these 36 correlations and variances with fewer terms? Yes, of course. The diagonal elements are all 1 and the off diagonal elements are symmetric. Thus, we have $n * (n - 1)$ correlations we want to model.

Eigen vectors of a 2 x 2 correlation matrix
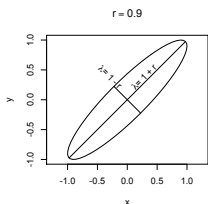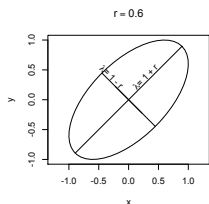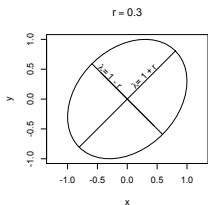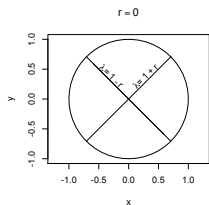
Although the length (eigen values) of the axes differ, their orientation (eigen vectors) are the same.

```
> r2 <- matrix(c(1,.6,.6,1),2,2)
> print(eigen(r2),2)

$values
[1] 1.6 0.4

$vectors
     [,1]  [,2]
[1,] 0.71 -0.71
[2,] 0.71  0.71
```

From eigen vectors to Principal Components

1. For n variables, there are n eigen vectors
   - There is no parsimony in thinking of the eigen vectors
   - Except that the vectors provide the orthogonal basis for the variables
2. Principal components are formed from the eigen vectors and eigen values
   - $\mathbf{R} = \mathbf{V}\lambda\mathbf{V}' = \mathbf{C}\mathbf{C}'$
   - $\mathbf{C} = \mathbf{V}\sqrt{\lambda}$
3. But there will still be as many Principal Components as variables, so what is the point?
4. Take just the first k Principal Components and see how well this reduced model fits the data.

The first principal component.

```
> pc1 <- principal(R,1)
> pc1

Uniquenesses:
   V1    V2    V3    V4    V5    V6
0.220 0.307 0.408 0.519 0.635 0.748
Loadings:
   PC1
V1 0.88
V2 0.83
V3 0.77
V4 0.69
V5 0.60
V6 0.50
               PC1
SS loadings    3.142
Proportion Var 0.524
```

```
#show  the model
> round(pc1$loadings %*% t(pc1$loadings),2)

   V1   V2   V3   V4   V5   V6
V1 0.77 0.73 0.68 0.61 0.53 0.44
V2 0.73 0.69 0.64 0.57 0.50 0.42
V3 0.68 0.64 0.59 0.53 0.46 0.38
V4 0.61 0.57 0.53 0.48 0.41 0.34
V5 0.53 0.50 0.46 0.41 0.36 0.30
V6 0.44 0.42 0.38 0.34 0.30 0.25

#find the residuals
> Rresid <- R - pc1$loadings %*% t(pc1$loadings)
> round(Rresid,2)

      V1    V2    V3    V4    V5    V6
V1  0.23 -0.01 -0.05 -0.07 -0.08 -0.08
V2 -0.01  0.31 -0.08 -0.09 -0.10 -0.09
V3 -0.05 -0.08  0.41 -0.11 -0.11 -0.10
V4 -0.07 -0.09 -0.11  0.52 -0.11 -0.10
V5 -0.08 -0.10 -0.11 -0.11  0.64 -0.10
V6 -0.08 -0.09 -0.10 -0.10 -0.10  0.75
```

The model fits pretty well, except that the diagonal is underestimated and the other correlations are over estimated.

Consider the following matrix

Correlations between 6 variables

Variable	V1	V2	V3	V4	V5	V6
V1	1.00	0.72	0.63	0.54	0.45	0.36
V2	0.72	1.00	0.56	0.48	0.40	0.32
V3	0.63	0.56	1.00	0.42	0.35	0.28
V4	0.54	0.48	0.42	1.00	0.30	0.24
V5	0.45	0.40	0.35	0.30	1.00	0.20
V6	0.36	0.32	0.28	0.24	0.20	1.00

Can we represent this in a simpler way?

$$R = FF' + U^2$$

or

$$R = CC'$$

Representing a correlation matrix with factors or components

Correlations between 6 variables

Variable	V1	V2	V3	V4	V5	V6
V1	1.00	0.72	0.63	0.54	0.45	0.36
V2	0.72	1.00	0.56	0.48	0.40	0.32
V3	0.63	0.56	1.00	0.42	0.35	0.28
V4	0.54	0.48	0.42	1.00	0.30	0.24
V5	0.45	0.40	0.35	0.30	1.00	0.20
V6	0.36	0.32	0.28	0.24	0.20	1.00

Table: $R = FF' + U^2$

Variable	loading
V1	0.9
V2	0.8
V3	0.7
V4	0.6
V5	0.5
V6	0.4

Table: $R = CC'$

Variable	
V1	0.88
V2	0.83
V3	0.77
V4	0.69
V5	0.60
V6	0.50

Factors vs. components

Originally developed by Spearman (1904) for the case of one common factor, and then later generalized by Thurstone (1947) and others to the case of multiple factors, factor analysis is probably the most frequently used and sometimes the most controversial psychometric procedure. The factor model, although seemingly very similar to the components model, is in fact very different. For rather than having components as linear sums of variables, in the factor model the variables are themselves linear sums of the unknown factors. That is, while components can be solved for by doing an *eigenvalue* or *singular value decomposition*, factors are estimated as best fitting solutions (Eckart & Young, 1936; Householder & Young, 1938), normally through iterative methods (Jöreskog, 1978; Lawley & Maxwell, 1963). Cattell (1965) referred to components analysis as a closed model and factor analysis as an open model, in that by explaining just the common variance, there was still more variance to explain.

Iterative principal axes factor analysis

Principal components represents a $n * n$ matrix in terms of the first
k components. It attempts to reproduce all of the **R** matrix.
Factor analysis on the other hand, attempts to model just the
common part of the matrix, which means all of the off-diagonal
elements and the common part of the diagonal (the
communalities). The non-common part, the *uniquenesses*, are
simply that which is left over. An easy to understand procedure is
principal axes factor analysis. This is similar to principal
components, except that it is done with a reduced matrix where
the diagonals are the communalities. The communalities can either
be specified a priori, estimated by such procedures as multiple
linear regression, or found by iteratively doing an eigenvalue
decomposition and repeatedly replacing the original 1s on the
diagonal with the the value of $1 - u^2$ where

$$\mathbf{U}^2 = diag(\mathbf{R} - \mathbf{FF}').$$

Principal axes as eigen values of a reduced matrix

That is, starting with the original correlation or covariance matrix, $\mathbf{R}$, find the k largest principal components, reproduce the matrix using those principal components. Find the resulting residual matrix, $\mathbf{R}^*$ and uniqueness matrix, $\mathbf{U}^2$ by

$$\mathbf{R}^* = \mathbf{R} - \mathbf{F}\mathbf{F}' \tag{1}$$

$$\mathbf{U}^2 = diag(\mathbf{R}^*)$$

and then, for iteration i, find $\mathbf{R}_i$ by replacing the diagonal of the original $\mathbf{R}$ matrix with 1 - diag($\mathbf{U}^2$) found on the previous step. Repeat this process until the change from one iteration to the next is arbitrarily small.

Simple Structure

The original solution of a principal components or principal axes factor analysis is a set of vectors that best account for the observed covariance or correlation matrix, and where the components or factors account for progressively less and less variance. But such a solution, although maximally efficient in describing the data, is rarely easy to interpret. But what makes a structure easy to interpret? Thurstone's answer, *simple structure*, consists of five rules (Thurstone, 1947, p 335):

> *(1) Each row of the oblique factor matrix* **V** *should have at least one zero.*
> *(2) For each column p of the factor matrix* **V** *there should be a distinct set of r linearly independent tests whose factor loadings v_{ip} are zero.*
> *(3) For every pair of columns of* **V** *there should be several tests whose entries v_{ip} vanish in one column but not in the other.*
> *(4) For every pair of columns of* **V**, *a large proportion of the tests should have zero entries in both columns. This applies to factor problems with four or five or more common factors.*
> *(5) For every pair of columns there should preferably be only a small number of tests with non-vanishing entries in both columns.*

Thurstone proposed to rotate the original solution to achieve simple structure.

Harman 8 physical measures

```
> data(Harman23.cor)
> lower.mat(Harman23.cor$cov)

               heght arm.s forrm lwr.l weght btr.d chst.g chst.w
height          1.00
arm.span        0.85  1.00
forearm         0.80  0.88  1.00
lower.leg       0.86  0.83  0.80  1.00
weight          0.47  0.38  0.38  0.44  1.00
bitro.diameter  0.40  0.33  0.32  0.33  0.76  1.00
chest.girth     0.30  0.28  0.24  0.33  0.73  0.58  1.00
chest.width     0.38  0.42  0.34  0.36  0.63  0.58  0.54   1.00
```

Preliminaries
00000000

The basic concepts
00

Principal Components
0000
000

Rotations and Transformations

References

Two solutions – loadings change, goodness of fits do not

```
> f2 <- fa(Harman23.cor$cov,2,rotate="none")
> f2
```

```
Factor Analysis using method =  minres
Call: fa(r = Harman23.cor$cov, nfactors = 2,
    rotate = "none")
Standardized loadings (pattern matrix)
                MR1   MR2   h2   u2
height          0.89 -0.19 0.83 0.17
arm.span        0.89 -0.31 0.89 0.11
forearm         0.86 -0.30 0.83 0.17
lower.leg       0.87 -0.22 0.80 0.20
weight          0.67  0.67 0.89 0.11
bitro.diameter  0.56  0.58 0.65 0.35
chest.girth     0.50  0.59 0.59 0.41
chest.width     0.56  0.40 0.47 0.53

                MR1  MR2
SS loadings     4.40 1.56
Proportion Var  0.55 0.19
Cumulative Var  0.55 0.74
Test of the hypothesis that 2 factors are sufficient.
The root mean square of the residuals (RMSR) is  0.02
The df corrected root mean square of the residuals
                is  0.03
Fit based upon off diagonal values = 1
```

```
> f2 <- fa(Harman23.cor$cov,2,rotate="varimax")
> f2
```

```
Factor Analysis using method =  minres
Call: fa(r = Harman23.cor$cov, nfactors = 2,
    rotate = "varimax")
Standardized loadings (pattern matrix)
                MR1  MR2   h2   u2
height          0.86 0.30 0.83 0.17
arm.span        0.92 0.20 0.89 0.11
forearm         0.89 0.19 0.83 0.17
lower.leg       0.86 0.26 0.80 0.20
weight          0.22 0.92 0.89 0.11
bitro.diameter  0.18 0.78 0.65 0.35
chest.girth     0.12 0.76 0.59 0.41
chest.width     0.27 0.63 0.47 0.53

                MR1  MR2
SS loadings     3.30 2.66
Proportion Var  0.41 0.33
Cumulative Var  0.41 0.74
```
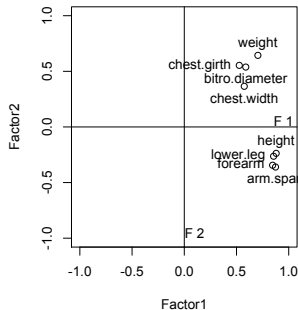
```
The root mean square of the residuals (RMSR)
    is  0.02
The df corrected root mean square of the residuals
    is  0.03
```
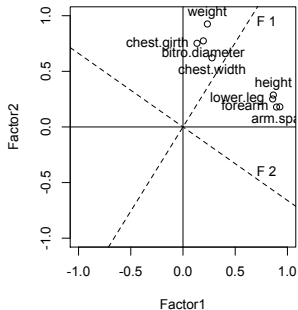
Alternative rotations

Oblique transfomations

Many of those who use factor analysis use it to identify theoretically meaningful constructs which they have no reason to believe are orthogonal. This has lead to the use of *oblique transformations* which allow the factors to be correlated. Although the term rotation is sometimes used for both *orthogonal* and *oblique* solutions, in the oblique case the factor matrix is not rotated so much as *transformed*.

Oblique transformations lead to the distinction between the *factor pattern* and *factor structure* matrices. The *factor pattern* matrix is the set of *regression weights* (loadings) from the latent factors to the observed variables. The *factor structure* matrix is the matrix of *correlations* between the factors and the observed variables. If the factors are uncorrelated, structure and pattern are identical. But, if the factors are correlated, the structure matrix ($\mathbf{S}$) is the pattern matrix ($\mathbf{F}$) times the factor intercorrelations ϕ
$\mathbf{S} = \mathbf{F}\phi <=> \mathbf{F} = \mathbf{S}\phi^{-1}$:

An oblique transformation of the Harman 8 physical variables

```
> f2t <- fa(Harman23.cor$cov,2,rotate="oblimin",n.obs=305)
> print(f2t)

Factor Analysis using method =  minres
Call: fa(r = Harman23.cor$cov, nfactors = 2, rotate = "oblimin", n.obs = 305)
               item   MR1   MR2   h2   u2
height            1  0.87  0.08 0.84 0.16
arm.span          2  0.96 -0.05 0.89 0.11
forearm           3  0.93 -0.04 0.83 0.17
lower.leg         4  0.88  0.04 0.81 0.19
weight            5  0.01  0.94 0.89 0.11
bitro.diameter    6  0.00  0.80 0.64 0.36
chest.girth       7 -0.06  0.79 0.59 0.41
chest.width       8  0.13  0.62 0.47 0.53

                MR1  MR2
SS loadings    3.37 2.58
Proportion Var 0.42 0.32
Cumulative Var 0.42 0.74

 With factor correlations of
     MR1  MR2
MR1 1.00 0.46
MR2 0.46 1.00
```

Preliminaries
00000000

The basic concepts
00

Principal Components
0000
000

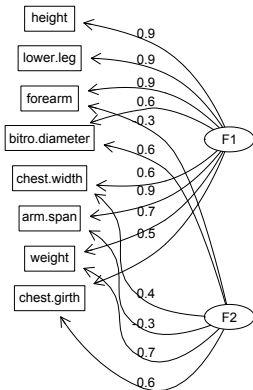Rotations and Transformations
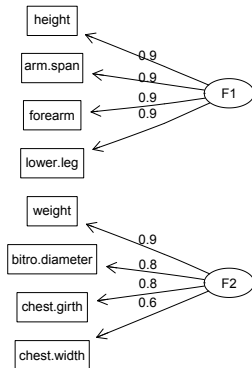
References
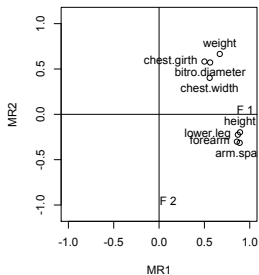
Oblique Transformations
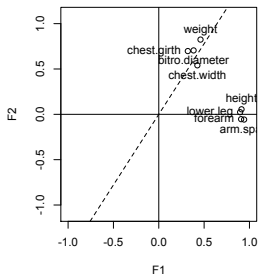
Another way to show simple structure

```
> simp24 <- sim.item(24,circum=FALSE)
> cor.plot(cor(simp24),main="A simple structure")
```
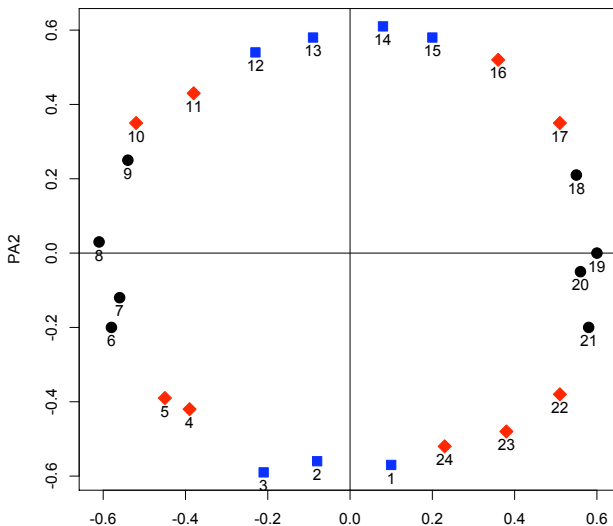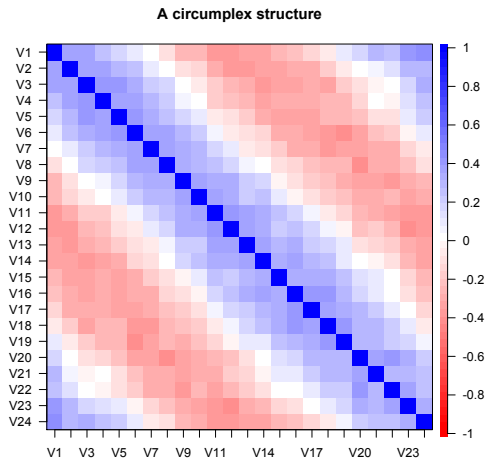
A circumplex is one alternative to simple structure

Circumplex structure

Another way of showing a circumplex – cor.plot

```
> circ24 <- sim.item(24,circum=TRUE)
> cor.plot(cor(circ24),main="A circumplex structure")
```

A circumplex structure

The Thurstone 9 variable problem – 9 measures of intellectual ability

```
> lower.mat(Thurstone)

                Sntnc Vcblr Snt.C Frs.L 4.L.W Sffxs Ltt.S Pdgrs Ltt.G
Sentences       1.00
Vocabulary      0.83  1.00
Sent.Completion 0.78  0.78  1.00
First.Letters   0.44  0.49  0.46  1.00
4.Letter.Words  0.43  0.46  0.42  0.67  1.00
Suffixes        0.45  0.49  0.44  0.59  0.54  1.00
Letter.Series   0.45  0.43  0.40  0.38  0.40  0.29  1.00
Pedigrees       0.54  0.54  0.53  0.35  0.37  0.32  0.56  1.00
Letter.Group    0.38  0.36  0.36  0.42  0.45  0.32  0.60  0.45  1.00
```

Three factors from Thurstone 9 variables

```
> f3 <- fa(Thurstone,3)
> f3


Factor Analysis using method =  minres
Call: fa(r = Thurstone, nfactors = 3)
Standardized loadings (pattern matrix) based upon correlation matrix
                  MR1   MR2   MR3   h2   u2
Sentences        0.91 -0.04  0.04 0.82 0.18
Vocabulary       0.89  0.06 -0.03 0.84 0.16
Sent.Completion  0.83  0.04  0.00 0.73 0.27
First.Letters    0.00  0.86  0.00 0.73 0.27
4.Letter.Words  -0.01  0.74  0.10 0.63 0.37
Suffixes         0.18  0.63 -0.08 0.50 0.50
Letter.Series    0.03 -0.01  0.84 0.72 0.28
Pedigrees        0.37 -0.05  0.47 0.50 0.50
Letter.Group    -0.06  0.21  0.64 0.53 0.47


                MR1  MR2  MR3
SS loadings    2.64 1.86 1.50
Proportion Var 0.29 0.21 0.17
Cumulative Var 0.29 0.50 0.67

 With factor correlations of
      MR1  MR2  MR3
MR1 1.00 0.59 0.54
MR2 0.59 1.00 0.52
MR3 0.54 0.52 1.00
```
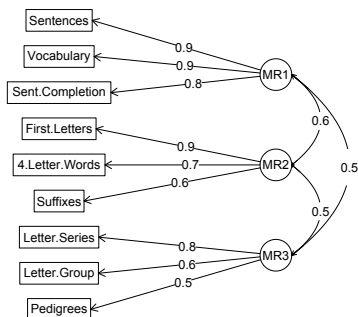
Preliminaries
○○○○○○○○

The basic concepts
○○

Principal Components
○○○○
○○○

Rotations and Transformations

References

A factor diagram of the Thurstone data

```
fa3 <-  fa(Thurstone,3,n.obs=213)
```
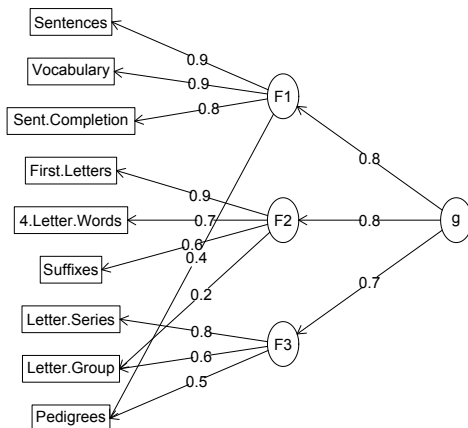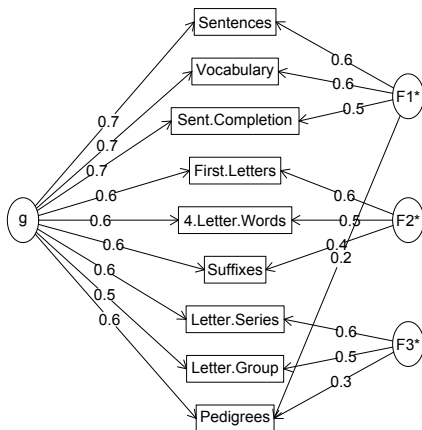

Factor Analysis

A hierarchical/multilevel solution to the Thurstone 9 variables

Hierarchical (multilevel) Structure

A bifactor solution using the Schmid Leiman transformation

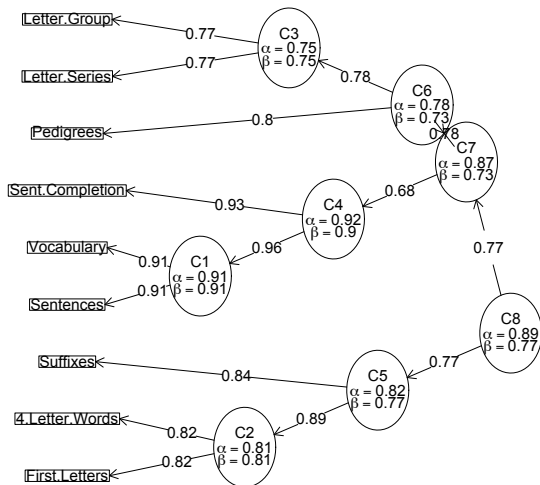**Omega with Schmid Leiman Transformation**

Hierarchical Clustering (e.g., `iclust`)

1. Find the proximity (e.g. correlation) matrix,
2. Identify the most similar pair of items
3. Combine this most similar pair of items to form a new variable (cluster),
4. Find the similarity of this cluster to all other items and clusters,
5. Repeat steps 2 and 3 until some criterion is reached (e.g., typically, if only one cluster remains or in `ICLUST` if there is a failure to increase reliability coefficients α or β).
6. Purify the solution by reassigning items to the most similar cluster center.

Yet another way to organize the data: cluster analysis
ICLUST

Preliminaries   The basic concepts   Principal Components   **Rotations and Transformations**   References
00000000   00   0000
                                      000

How many factors – no right answer, one wrong answer

1. Statistical
   - Extracting factors until the χ^2 of the residual matrix is not significant.
   - Extracting factors until the change in χ^2 from factor n to factor n+1 is not significant.
2. Rules of Thumb
   - Parallel Extracting factors until the eigenvalues of the real data are less than the corresponding eigenvalues of a random data set of the same size (*parallel analysis*)
   - Plotting the magnitude of the successive eigenvalues and applying the *scree test*.
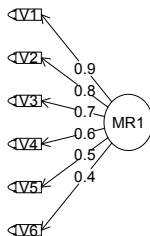3. Interpretability
   - Extracting factors as long as they are interpretable.
   - Using the *Very Simple Structure* Criterion (VSS)
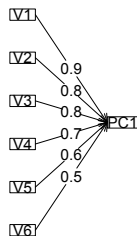   - Using the Minimum Average Partial criterion (MAP).
4. Eigen Value of 1 rule

Preliminaries
○○○○○○○○

The basic concepts
○○

Principal Components
○○○○
○○○

Rotations and Transformations

References

The factor model versus the components model

Factor model

Components model

Cattell, R. B. (1965). A biometrics invited paper. factor analysis: An introduction to essentials i. the purpose and underlying models. *Biometrics*, *21*(1), 190–215.

Eckart, C. & Young, G. (1936). The approximation of one matrix by another of lower rank. *Psychometrika*, *1*(3), 211–218.

Householder, A. S. & Young, G. (1938). Matrix approximation and latent roots. *The American Mathematical Monthly*, *45*(3), 165–171.

Jöreskog, K. (1978). Structural analysis of covariance and correlation matrices. *Psychometrika*, *43*(4), 443–477.

Lawley, D. N. & Maxwell, A. E. (1963). *Factor analysis as a statistical method*. London: Butterworths.

Spearman, C. (1904). "General Intelligence," objectively determined and measured. *American Journal of Psychology*, *15*(2), 201–292.

Thurstone, L. L. (1947). *Multiple-factor analysis: a development and expansion of The vectors of the mind*. Chicago, Ill.: The University of Chicago Press.