Psychology 454: Latent Variable Modeling
Week 4: Latent models of real data

William Revelle

Department of Psychology
Northwestern University
Evanston, Illinois USA

January, 2011

Outline

Real data rather than simulated

- Prior examples were all artificial data sets.
    - These have the advantage that we know truth.
    - They have the disadvantage that they don't give us experience with real data.
- Some classic data sets are available in *psych*, *lavaan* and *sem*.
- Using the `data` function to see what is available.
    - data(package="psych")

Some of the psych data sets

data(package="psych")

```
Bechtoldt.1 (bifactor)        Seven data sets showing a bifactor solution.
Bechtoldt.2 (bifactor)        Seven data sets showing a bifactor solution.
Chen (Schmid)                 12 variables created by Schmid and Leiman to show the
                              Schmid-Leiman Transformation
Harman.Burt (Harman)          Two data sets from Harman (1967). 9 cognitive variables from
                              Holzinger and 8 emotional variables from Burt
Harman.Holzinger (Harman)     Two data sets from Harman (1967). 9 cognitive variables from
                              Holzinger and 8 emotional variables from Burt
Holzinger (bifactor)          Seven data sets showing a bifactor solution.
Holzinger.9 (bifactor)        Seven data sets showing a bifactor solution.
Reise (bifactor)              Seven data sets showing a bifactor solution.
Schmid                        12 variables created by Schmid and Leiman to show the
                              Schmid-Leiman Transformation
Thurstone (bifactor)          Seven data sets showing a bifactor solution.
Thurstone.33 (bifactor)       Seven data sets showing a bifactor solution.
Tucker                        9 Cognitive variables discussed by Tucker and Lewis (1973)
West (Schmid)                 12 variables created by Schmid and Leiman to show the
                              Schmid-Leiman Transformation
all.income (income)           US family income from US census 2008
bfi                           25 Personality items representing 5 factors
blot                          Bond's Logical Operations Test - BLOT
bock.table (bock)             Bock and Liberman (1970) data set of 1000 observations of the
                              LSAT
burt                          11 emotional variables from Burt (1915)
cities                        Distances between 11 US cities
city.location (cities)        Distances between 11 US cities
cubits                        Galton's example of the relationship between height and 'cubit'
                              or forearm length
epi.bfi                       13 personality scales from the Eysenck Personality Inventory
                              and Big 5 inventory
flat (affect)                 Two data sets of affect and arousal scores as a function of
```

Also can read in from clipboard, web, from hard drive and import

- To read from clipboard: `read.clipboard()`
  - read.clipboard(header = TRUE, ...) #assumes headers and tab or space delimited
  - read.clipboard.csv(header=TRUE,sep=',',...) #assumes headers and comma delimited
  - read.clipboard.tab(header=TRUE,sep='\t',...) #assumes headers and tab delimited
  - read.clipboard.lower(diag=TRUE,names=FALSE,...) #read in a matrix given the lower off diagonal
  - read.clipboard.upper(diag=TRUE,names=FALSE,...)
  - read.clipboard.fwf(header=FALSE,widths=rep(1,10),...) #read in data using a fixed format width (see read.fwf for instructions)
- To read from a file
  - fn <- file.choose() #this opens the system directory – navigate it to your file
  - my.data <- read.table(fn)

More on getting your data

```
#specify the name and address of the remote file
datafilename <- "http://personality-project.org/r/datasets/maps.mixx.epi.bfi.dat
#datafilename <- "Desktop/epi.big5.txt"  #read from local directory  or
# datafilename <- file.choose()     # use the OS to find the file
#in all cases
person.data  <- read.table(datafilename,header=TRUE)  #read the data file

#Alternatively, to read in a comma delimited file:
#person.data  <- read.table(datafilename,header=TRUE,sep=",")


names(person.data)  #list the names of the variables

library(foreign) #allows you to read SPSS files   e.g.,
datafilename <- "/Users/bill/Library/Favorites/R.tutorial/datasets/finkel.sav"
eli.data <- read.spss(datafilename, use.value.labels=TRUE, to.data.frame=TRUE)
```

Use the bifactor examples

The study of latent variable models in general and sem in particular is the combination of measurement models with structural models. The bifactor examples are useful to understand issues in reliability and estimating measurement models.

```
> data(bifactor)
> ?bifactor
```

```
Two more data sets with similar structures are found in the Harman data set.

Bechtoldt.1: 17 x 17 correlation matrix of ability tests, N = 212.
Bechtoldt.2: 17 x 17 correlation matrix of ability tests, N = 213.
Holzinger: 14 x 14 correlation matrix of ability tests, N = 355
Holzinger.9: 9 x 9 correlation matrix of ability tests, N = 145
Reise: 16 x 16 correlation matrix of health satisfaction items. N = 35,000
Thurstone: 9 x 9 correlation matrix of ability tests, N = 213
Thurstone.33: Another 9 x 9 correlation matrix of ability items, N=4175
```

Thurstone correlation matrix

```
> colnames(Thurstone) <- paste("V",1:9,sep="")
> Thurstone
```

	V1	V2	V3	V4	V5	V6	V7	V8	V9
Sentences	1.000	0.828	0.776	0.439	0.432	0.447	0.447	0.541	0.380
Vocabulary	0.828	1.000	0.779	0.493	0.464	0.489	0.432	0.537	0.358
Sent.Completion	0.776	0.779	1.000	0.460	0.425	0.443	0.401	0.534	0.359
First.Letters	0.439	0.493	0.460	1.000	0.674	0.590	0.381	0.350	0.424
4.Letter.Words	0.432	0.464	0.425	0.674	1.000	0.541	0.402	0.367	0.446
Suffixes	0.447	0.489	0.443	0.590	0.541	1.000	0.288	0.320	0.325
Letter.Series	0.447	0.432	0.401	0.381	0.402	0.288	1.000	0.555	0.598
Pedigrees	0.541	0.537	0.534	0.350	0.367	0.320	0.555	1.000	0.452
Letter.Group	0.380	0.358	0.359	0.424	0.446	0.325	0.598	0.452	1.000

The number of factors problem

- "The number of factors problem is easy, I solve it everyday before breakfast. The problem is the right number" Kaiser.
  - Extract factors until χ^2 is not significant.
  - Extract factors until change in χ^2 is not significant.
  - Parallel analysis.
  - scree test
  - Very Simple Structure (VSS)
  - Minimum Absolute Partial (Velicer's MAP test)

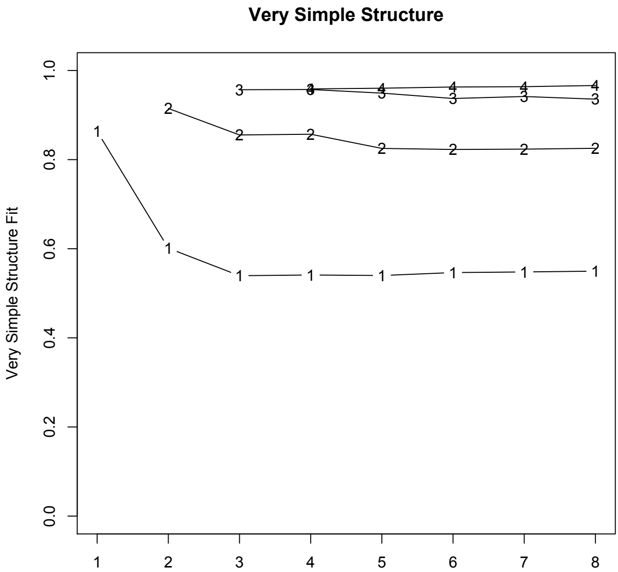
Number of factors?

```
> vss <- VSS(Thurstone,n.obs=213,SMC=FALSE)
> vss


Very Simple Structure
Call: VSS(x = Thurstone, n.obs = 213, SMC = FALSE)
VSS complexity 1 achieves a maximimum of 0.86  with  1  factors
VSS complexity 2 achieves a maximimum of 0.91  with  2  factors

The Velicer MAP criterion achieves a minimum of 0.07  with  3  factors

Velicer MAP
[1] 0.07 0.07 0.07 0.11 0.20 0.31 0.59 1.00

Very Simple Structure Complexity 1
[1] 0.86 0.60 0.54 0.54 0.54 0.55 0.55 0.55

Very Simple Structure Complexity 2
[1] 0.00 0.91 0.86 0.86 0.83 0.82 0.82 0.83
```

χ^2

```
> print(vss$vss.stats[,1:3],digits=3)

  dof    chisq     prob
1  27 2.32e+02 2.14e-34
2  19 8.30e+01 5.56e-10
3  12 2.82e+00 9.97e-01
4   6 1.49e+00 9.60e-01
5   1 2.44e-01 6.21e-01
6  -3 4.80e-07       NA
7  -6 4.16e-09       NA
8  -8 0.00e+00       NA
```

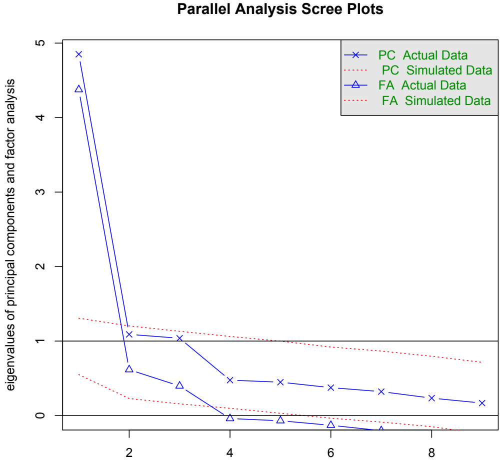
A Very Simple Structure Plot



Parallel analysis

> fa.parallel(Thurstone,n.obs=213)

Parallel analysis suggests that the number of factors = 3 and the number of components = 1



Minimal Residual FA

fa(Thurstone,n.obs=213)

```
> fa(Thurstone,n.obs=213)
Factor Analysis using method =  minres
Call: fa(r = Thurstone, n.obs = 213)
Standardized loadings based upon correlation matrix
    MR1   h2   u2
V1 0.87 0.75 0.25
V2 0.88 0.77 0.23
V3 0.83 0.70 0.30
V4 0.62 0.39 0.61
V5 0.61 0.37 0.63
V6 0.59 0.34 0.66
V7 0.57 0.32 0.68
V8 0.64 0.41 0.59
V9 0.52 0.27 0.73

                MR1
SS loadings    4.32
Proportion Var 0.48

Test of the hypothesis that 1 factor is sufficient.

The degrees of freedom for the null model are  36  and the objective function w
```

Maximum likelihood

```
> fa(Thurstone,n.obs=213,fm="mle")
Factor Analysis using method =  ml
Call: fa(r = Thurstone, n.obs = 213, fm = "mle")
Standardized loadings based upon correlation matrix
    ML1   h2   u2
V1 0.88 0.78 0.22
V2 0.90 0.80 0.20
V3 0.85 0.72 0.28
V4 0.59 0.35 0.65
V5 0.57 0.33 0.67
V6 0.57 0.32 0.68
V7 0.54 0.29 0.71
V8 0.63 0.40 0.60
V9 0.49 0.24 0.76

                ML1
SS loadings    4.22
Proportion Var 0.47

Test of the hypothesis that 1 factor is sufficient.

The degrees of freedom for the null model are  36  and the objective function w
```

Comparing min res and mle

```
Standardized loadings based upon correlation matrix
    MR1   h2   u2
V1 0.87 0.75 0.25
V2 0.88 0.77 0.23
V3 0.83 0.70 0.30
V4 0.62 0.39 0.61
V5 0.61 0.37 0.63
V6 0.59 0.34 0.66
V7 0.57 0.32 0.68
V8 0.64 0.41 0.59
V9 0.52 0.27 0.73

                MR1
SS loadings    4.32
Proportion Var 0.48
Chi Square =  231.55
     with prob <  2.1e-34
```

```
    ML1   h2   u2
V1 0.88 0.78 0.22
V2 0.90 0.80 0.20
V3 0.85 0.72 0.28
V4 0.59 0.35 0.65
V5 0.57 0.33 0.67
V6 0.57 0.32 0.68
V7 0.54 0.29 0.71
V8 0.63 0.40 0.60
V9 0.49 0.24 0.76

                ML1
SS loadings    4.22
Proportion Var 0.47
Chi Square =  228.59
     with prob <  8e-34
```

2 maximum likelihood factors

```
> f2 <- fa(Thurstone,2,n.obs=213,fm="mle")
> f2
Factor Analysis using method =  ml
Call: fa(r = Thurstone, nfactors = 2, n.obs = 213, fm = "mle")
Standardized loadings based upon correlation matrix
     ML1   ML2   h2   u2
V1  0.94 -0.05 0.83 0.17
V2  0.89  0.03 0.82 0.18
V3  0.85  0.02 0.73 0.27
V4 -0.02  0.84 0.68 0.32
V5 -0.04  0.84 0.66 0.34
V6  0.13  0.59 0.46 0.54
V7  0.28  0.35 0.32 0.68
V8  0.50  0.17 0.39 0.61
V9  0.12  0.49 0.32 0.68
                 ML1  ML2
SS loadings     2.92 2.29
Proportion Var 0.32 0.25
Cumulative Var 0.32 0.58
 With factor correlations of
     ML1  ML2
ML1 1.00 0.64
ML2 0.64 1.00
```

2 factor goodness of fits

```
Test of the hypothesis that 2 factors are sufficient.

The degrees of freedom for the null model are  36  and the objective function
            was  5.2 with Chi Square of  1081.97
The degrees of freedom for the model are 19  and the
             objective function was  0.4

The root mean square of the residuals is  0.05
The df corrected root mean square of the residuals is  0.1
The number of observations was  213  with Chi Square =  82.84
                       with prob <  6e-10

Tucker Lewis Index of factoring reliability =  0.884
RMSEA index =  0.128  and the 90 % confidence intervals are  0.127 0.131
BIC =  -19.02
Fit based upon off diagonal values = 0.98
Measures of factor score adequacy
                                                ML1  ML2
Correlation of scores with factors             0.97 0.93
Multiple R square of scores with factors       0.93 0.86
Minimum correlation of possible factor scores  0.86 0.72
```

3 MLE factors of Thurstone data

```
> f3 <- fa(Thurstone,3,n.obs=213,fm="mle")
> f3
Call: fa(r = Thurstone, nfactors = 3, n.obs = 213, fm = "mle")
Standardized loadings based upon correlation matrix
     ML1   ML2   ML3   h2   u2
V1  0.91 -0.04  0.04 0.83 0.17
V2  0.89  0.06 -0.03 0.84 0.16
V3  0.83  0.04  0.00 0.73 0.27
V4  0.00  0.86  0.01 0.73 0.27
V5 -0.01  0.74  0.10 0.63 0.37
V6  0.18  0.63 -0.08 0.50 0.50
V7  0.03 -0.01  0.84 0.72 0.28
V8  0.37 -0.05  0.47 0.50 0.50
V9 -0.06  0.21  0.64 0.53 0.47
                 ML1  ML2  ML3
SS loadings     2.64 1.86 1.49
Proportion Var 0.29 0.21 0.17
Cumulative Var 0.29 0.50 0.67
 With factor correlations of
     ML1  ML2  ML3
ML1 1.00 0.59 0.54
ML2 0.59 1.00 0.52
ML3 0.54 0.52 1.00
```

with fit statistics

```
Test of the hypothesis that 3 factors are sufficient.

The degrees of freedom for the null model are  36  and the
         objective function was  5.2 with Chi Square of  1081.97
The degrees of freedom for the model are 12  and the
         objective function was  0.01

The root mean square of the residuals is  0
The df corrected root mean square of the residuals is  0.01
The number of observations was  213  with
            Chi Square =  2.82  with prob <  1

Tucker Lewis Index of factoring reliability =  1.027
RMSEA index =  0  and the 90 % confidence intervals are  0 0.023
BIC =  -61.51
Fit based upon off diagonal values = 1
Measures of factor score adequacy
                                                ML1  ML2  ML3
Correlation of scores with factors             0.96 0.92 0.90
Multiple R square of scores with factors       0.93 0.85 0.81
Minimum correlation of possible factor scores  0.86 0.71 0.63
```

Orthogonal Rotations and oblique Transformations

- Should the latent variables be allowed to correlate?
  - Factors as extracted are orthogonal.
  - Factors as interpreted typically are allowed to be oblique
- Consider five cases
  - unrotated: factors as extracted (rotate="none")
  - rotated using VARIMAX (or Quartimax)
  - transformed using PROCRUSTES
  - transformed using oblimin
  - transformed using geominQ

Unrotated versus Varimax Rotation

```
> fa(Thurstone,3,n.obs=213,rotate="none")

Factor Analysis using method =  minres
Call: fa(r = Thurstone, nfactors = 3, n.obs = 213, rotate = "none")
Standardized loadings based upon correlation matrix
    MR1   MR2   MR3   h2   u2
V1 0.87 -0.27  0.02 0.82 0.18
V2 0.88 -0.24 -0.06 0.84 0.16
V3 0.83 -0.22 -0.03 0.73 0.27
V4 0.66  0.45 -0.32 0.73 0.27
V5 0.63  0.43 -0.21 0.63 0.37
V6 0.60  0.24 -0.29 0.50 0.50
V7 0.60  0.32  0.50 0.72 0.28
V8 0.65  0.05  0.29 0.50 0.50
V9 0.54  0.38  0.30 0.53 0.47

                MR1  MR2  MR3
SS loadings    4.47 0.86 0.67
Proportion Var 0.50 0.10 0.07
Cumulative Var 0.50 0.59 0.67
```

```
> fa(Thurstone,3,n.obs=213,rotate="Varimax")

    MR1  MR2  MR3   h2   u2
V1 0.86 0.20 0.22 0.82 0.18
V2 0.85 0.27 0.18 0.84 0.16
V3 0.80 0.24 0.19 0.73 0.27
V4 0.29 0.78 0.20 0.73 0.27
V5 0.27 0.70 0.26 0.63 0.37
V6 0.36 0.60 0.10 0.50 0.50
V7 0.28 0.18 0.78 0.72 0.28
V8 0.48 0.15 0.50 0.50 0.50
V9 0.20 0.32 0.62 0.53 0.47

                MR1  MR2  MR3
SS loadings    2.73 1.78 1.48
Proportion Var 0.30 0.20 0.16
Cumulative Var 0.30 0.50 0.67
```

Varimax versus Quartimax

```
> fa(Thurstone,3,n.obs=213,rotate="Varimax")

    MR1  MR2  MR3   h2   u2
V1 0.86 0.20 0.22 0.82 0.18
V2 0.85 0.27 0.18 0.84 0.16
V3 0.80 0.24 0.19 0.73 0.27
V4 0.29 0.78 0.20 0.73 0.27
V5 0.27 0.70 0.26 0.63 0.37
V6 0.36 0.60 0.10 0.50 0.50
V7 0.28 0.18 0.78 0.72 0.28
V8 0.48 0.15 0.50 0.50 0.50
V9 0.20 0.32 0.62 0.53 0.47

                MR1  MR2  MR3
SS loadings    2.73 1.78 1.48
Proportion Var 0.30 0.20 0.16
Cumulative Var 0.30 0.50 0.67
```

```
> fa(Thurstone,3,n.obs=213,rotate="quartimax")

    MR1  MR2  MR3   h2   u2
V1 0.91 0.02 0.05 0.82 0.18
V2 0.91 0.09 0.01 0.84 0.16
V3 0.85 0.07 0.03 0.73 0.27
V4 0.47 0.71 0.11 0.73 0.27
V5 0.45 0.63 0.18 0.63 0.37
V6 0.48 0.51 0.01 0.50 0.50
V7 0.45 0.12 0.71 0.72 0.28
V8 0.58 0.05 0.40 0.50 0.50
V9 0.37 0.27 0.56 0.53 0.47

                MR1  MR2  MR3
SS loadings    3.70 1.27 1.03
Proportion Var 0.41 0.14 0.11
Cumulative Var 0.41 0.55 0.67
```

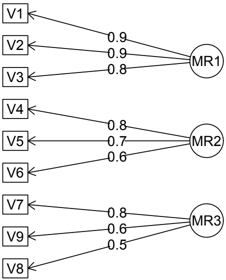
Varimax versus Quartimin

```
> fa(Thurstone,3,n.obs=213,
       rotate="Varimax")

    MR1  MR2  MR3  h2   u2
V1 0.86 0.20 0.22 0.82 0.18
V2 0.85 0.27 0.18 0.84 0.16
V3 0.80 0.24 0.19 0.73 0.27
V4 0.29 0.78 0.20 0.73 0.27
V5 0.27 0.70 0.26 0.63 0.37
V6 0.36 0.60 0.10 0.50 0.50
V7 0.28 0.18 0.78 0.72 0.28
V8 0.48 0.15 0.50 0.50 0.50
V9 0.20 0.32 0.62 0.53 0.47

                MR1  MR2  MR3
SS loadings     2.73 1.78 1.48
Proportion Var  0.30 0.20 0.16
Cumulative Var  0.30 0.50 0.67
```

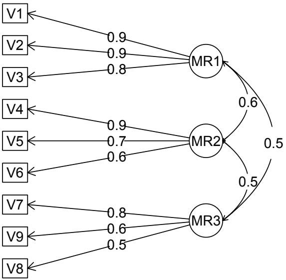
```
> fa(Thurstone,3,n.obs=213,rotate="quartimin")

     MR1   MR2   MR3  h2   u2
V1  0.91 -0.04  0.04 0.82 0.18
V2  0.89  0.06 -0.03 0.84 0.16
V3  0.83  0.04  0.00 0.73 0.27
V4  0.00  0.86  0.00 0.73 0.27
V5 -0.01  0.74  0.10 0.63 0.37
V6  0.18  0.63 -0.08 0.50 0.50
V7  0.03 -0.01  0.84 0.72 0.28
V8  0.37 -0.05  0.47 0.50 0.50
V9 -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
```

Visualizing the difference: Varimax versus quartimin

Orthogonal Rotation



Oblique Transformation



Specifying the parameters in sem using the RAM notation

- Need to specify each path
- Need to specify the error variance paths as well
  - A short cut can be done by using *psych*
  - See the vignette ‘`psych for sem`‘” for more details
- Apply this to the Thurstone problem

create the sem commands by using psych

```
f3 <- fa(Thurstone,3,fm='mle')
mod3 <- structure.diagram(f3,cut=.45,errors=TRUE)
mod3
```

```
      Path         Parameter Value
 [1,] "ML1->V1"    "F1V1"    NA
 [2,] "ML1->V2"    "F1V2"    NA
 [3,] "ML1->V3"    "F1V3"    NA
 [4,] "ML2->V4"    "F2V4"    NA
 [5,] "ML2->V5"    "F2V5"    NA
 [6,] "ML2->V6"    "F2V6"    NA
 [7,] "ML3->V7"    "F3V7"    NA
 [8,] "ML3->V8"    "F3V8"    NA
 [9,] "ML3->V9"    "F3V9"    NA
[10,] "V1<->V1"    "x1e"     NA
[11,] "V2<->V2"    "x2e"     NA
...
[18,] "V9<->V9"    "x9e"     NA
[19,] "ML2<->ML1" "rF2F1"    NA
[20,] "ML3<->ML1" "rF3F1"    NA
[21,] "ML3<->ML2" "rF3F2"    NA
[22,] "ML1<->ML1" NA         "1"
[23,] "ML2<->ML2" NA         "1"
[24,] "ML3<->ML3" NA         "1"
```

Running sem

```
> rownames(Thurstone) <- colnames(Thurstone) #to get the names to match the modl
> sem3 <- sem(mod3,Thurstone,N=213)
> summary(sem3,digits=2)
```

```
 Model Chisquare =  38   Df =  24 Pr(>Chisq) = 0.033
 Chisquare (null model) =  1102   Df =  36
 Goodness-of-fit index =  0.96
 Adjusted goodness-of-fit index =  0.92
 RMSEA index =  0.053   90% CI: (0.015, 0.083)
 Bentler-Bonnett NFI =  0.97
 Tucker-Lewis NNFI =  0.98
 Bentler CFI =  0.99
 SRMR =  0.044
 BIC =  -90

 Normalized Residuals
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  -0.97   -0.42    0.00    0.04    0.09    1.63
```

With parameter estimates

```
 Parameter Estimates
      Estimate Std Error z value Pr(>|z|)
F1V1  0.90     0.054     16.7    0.0e+00  V1 <--- ML1
F1V2  0.91     0.054     17.0    0.0e+00  V2 <--- ML1
F1V3  0.86     0.056     15.3    0.0e+00  V3 <--- ML1
F2V4  0.84     0.061     13.8    0.0e+00  V4 <--- ML2
F2V5  0.80     0.062     12.9    0.0e+00  V5 <--- ML2
F2V6  0.70     0.064     10.9    0.0e+00  V6 <--- ML2
F3V7  0.78     0.065     12.0    0.0e+00  V7 <--- ML3
F3V8  0.72     0.067     10.7    0.0e+00  V8 <--- ML3
F3V9  0.70     0.067     10.5    0.0e+00  V9 <--- ML3
x1e   0.18     0.028      6.4    1.7e-10  V1 <--> V1
x2e   0.16     0.028      5.9    3.0e-09  V2 <--> V2
x3e   0.27     0.033      8.0    1.6e-15  V3 <--> V3
x4e   0.30     0.051      5.9    2.7e-09  V4 <--> V4
x5e   0.36     0.052      7.0    3.4e-12  V5 <--> V5
x6e   0.51     0.060      8.4    0.0e+00  V6 <--> V6
x7e   0.39     0.062      6.3    2.3e-10  V7 <--> V7
x8e   0.48     0.065      7.4    1.8e-13  V8 <--> V8
x9e   0.51     0.065      7.7    9.5e-15  V9 <--> V9
rF2F1 0.64     0.051     12.6    0.0e+00  ML1 <--> ML2
rF3F1 0.67     0.054     12.5    0.0e+00  ML1 <--> ML3
rF3F2 0.64     0.059     10.7    0.0e+00  ML2 <--> ML3
```

Exploratory bifactor model

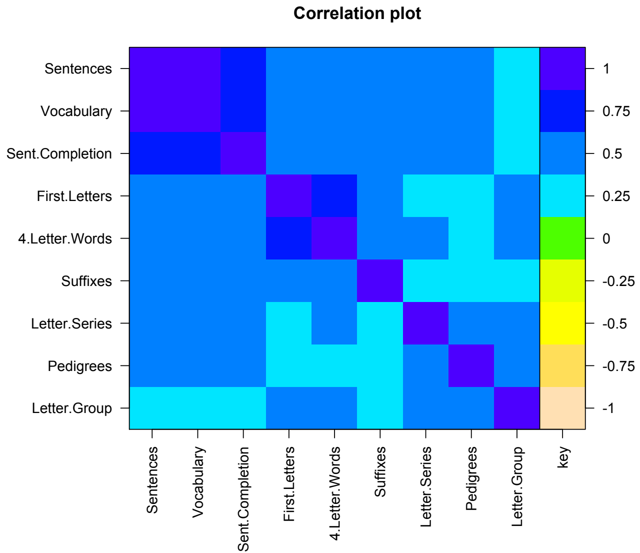
```
omt <- omega(Thurstone)
omega.diagram(omt,sl=FALSE)
```

```
Omega
Call: omega(m = Thurstone)
Alpha:                 0.89
G.6:                   0.91
Omega Hierarchical:    0.74
Omega H asymptotic:    0.79
Omega Total            0.93

Schmid Leiman Factor loadings greater than  0.2
      g  F1*  F2*  F3*   h2   u2   p2
V1 0.71 0.57            0.82 0.18 0.61
V2 0.73 0.55            0.84 0.16 0.63
V3 0.68 0.52            0.73 0.27 0.63
V4 0.65      0.56       0.73 0.27 0.57
V5 0.62      0.49       0.63 0.37 0.61
V6 0.56      0.41       0.50 0.50 0.63
V7 0.59           0.61 0.72 0.28 0.48
V8 0.58 0.23      0.34 0.50 0.50 0.66
V9 0.54           0.46 0.53 0.47 0.56
```

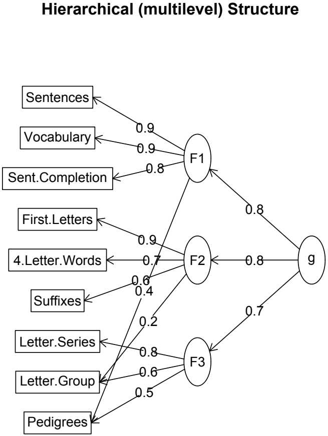
With fit statistics

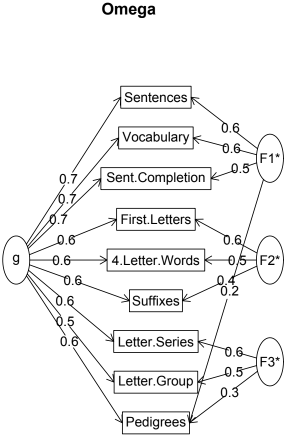
```
With eigenvalues of:
   g  F1*  F2*  F3*
3.58 0.96 0.74 0.71
general/max  3.71   max/min =   1.35
mean percent general =  0.6    with sd =  0.05 and cv of  0.09
The degrees of freedom are 12  and the fit is  0.01
The number of observations was  213  with Chi Square =  2.82  with prob <  1
The root mean square of the residuals is  0
The df corrected root mean square of the residuals is  0.01
RMSEA index =  0  and the 90 % confidence intervals are  0 0.023
BIC =  -61.51
Compare this with the adequacy of just a general factor and no group factors
The degrees of freedom for just the general factor are 27  and the fit is  1.48
The number of observations was  213  with Chi Square =  307.1  with prob <  2.8
The root mean square of the residuals is  0.1
The df corrected root mean square of the residuals is  0.16
RMSEA index =  0.224  and the 90 % confidence intervals are  0.223 0.226
BIC =  162.35
Measures of factor score adequacy
                                                 g  F1*  F2*  F3*
Correlation of scores with factors            0.86 0.73 0.72 0.75
Multiple R square of scores with factors      0.74 0.54 0.52 0.56
Minimum correlation of factor score estimates 0.49 0.08 0.03 0.11
```

A correlation plot to show a general factor



Two ways of showing a hierarchical structure





Exploratory versus confirmatory bifactor model

- `omega` will do exploratory bifactoring
- `omegaSem` will do exploratory and then confirmatory based upon that solution.
  - This is not true "confirmatory" in that the solution was decided on in an exploratory fashion.
  - True confirmatory tests a prior hypothesis
- `omegaSem` reports both exploratory and confirmatory results.
- The sem model is hidden as part of the structure but may be found from the `str` command

omegaSem

```
Alpha:                 0.89
G.6:                   0.91
Omega Hierarchical:    0.74
Omega H asymptotic:    0.79
Omega Total            0.93

Schmid Leiman Factor loadings greater than  0.2
                   g   F1*   F2*   F3*   h2   u2   p2
Sentences       0.71  0.57              0.82 0.18 0.61
Vocabulary      0.73  0.55              0.84 0.16 0.63
Sent.Completion 0.68  0.52              0.73 0.27 0.63
First.Letters   0.65        0.56        0.73 0.27 0.57
4.Letter.Words  0.62        0.49        0.63 0.37 0.61
Suffixes        0.56        0.41        0.50 0.50 0.63
Letter.Series   0.59              0.61 0.72 0.28 0.48
Pedigrees       0.58  0.23        0.34 0.50 0.50 0.66
Letter.Group    0.54              0.46 0.53 0.47 0.56

With eigenvalues of:
   g  F1*  F2*  F3*
3.58 0.96 0.74 0.71
```

```
 Omega Hierarchical from a confirmatory model
      using sem =  0.79
 Omega Total   from a confirmatory model
      using sem =  0.93
With loadings of
                   g  F1*  F2*  F3*   h2   u2
Sentences       0.77 0.49            0.83 0.17
Vocabulary      0.79 0.45            0.83 0.17
Sent.Completion 0.75 0.40            0.73 0.27
First.Letters   0.61      0.61       0.75 0.25
4.Letter.Words  0.60      0.51       0.61 0.39
Suffixes        0.57      0.39       0.48 0.52
Letter.Series   0.57           0.73 0.85 0.15
Pedigrees       0.66           0.25 0.50 0.50
Letter.Group    0.53           0.41 0.45 0.55

With eigenvalues of:
   g  F1*  F2*  F3*
3.88 0.61 0.79 0.76
```

The sem model is an object in omegaSem output

```
omts <- omegaSem(Thurstone,n.obs=213)
omts$omegaSem$model

    Path                              Parameter           Initial Value
 [1,] "g->Sentences"                    "Sentences"          NA
 [2,] "g->Vocabulary"                   "Vocabulary"         NA
 [...
 [9,] "g->Letter.Group"                 "Letter.Group"       NA
[10,] "F1*->Sentences"                  "F1*Sentences"       NA
[11,] "F1*->Vocabulary"                 "F1*Vocabulary"      NA
[12,] "F1*->Sent.Completion"            "F1*Sent.Completion" NA
[13,] "F2*->First.Letters"              "F2*First.Letters"   NA
[14,] "F2*->4.Letter.Words"             "F2*4.Letter.Words"  NA
[15,] "F2*->Suffixes"                   "F2*Suffixes"        NA
[16,] "F3*->Letter.Series"              "F3*Letter.Series"   NA
[17,] "F3*->Pedigrees"                  "F3*Pedigrees"       NA
[18,] "F3*->Letter.Group"               "F3*Letter.Group"    NA
[19,] "Sentences<->Sentences"           "e1"                 NA
[20,] "Vocabulary<->Vocabulary"         "e2"                 NA
...
[27,] "Letter.Group<->Letter.Group"     "e9"                 NA
[28,] "F1*<->F1*"                       NA                   "1"
[29,] "F2*<->F2*"                       NA                   "1"
[30,] "F3*<->F3*"                       NA                   "1"
[31,] "g <->g"                          NA                   "1"
```

Using the sem model from omeaSem

```
> sem.t.om <- sem(omts$omegaSem$model,Thurstone,N=213)
> summary(sem.t.om)
```

```
 Model Chisquare =  24.216   Df =  18 Pr(>Chisq) = 0.14807
 Chisquare (null model) =  1101.9   Df =  36
 Goodness-of-fit index =  0.97578
 Adjusted goodness-of-fit index =  0.93944
 RMSEA index =  0.040361   90% CI: (NA, 0.077994)
 Bentler-Bonnett NFI =  0.97802
 Tucker-Lewis NNFI =  0.98834
 Bentler CFI =  0.99417
 SRMR =  0.034895
 BIC =  -72.287

 Normalized Residuals
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-0.8210 -0.3340  0.0000  0.0282  0.1560  1.8000
```

sem parameter estimates

```
Parameter Estimates
                  Estimate Std Error z value  Pr(>|z|)
Sentences         0.76787  0.072626  10.57291 0.0000e+00 Sentences <--- g
Vocabulary        0.79092  0.072418  10.92170 0.0000e+00 Vocabulary <--- g
Sent.Completion   0.75362  0.073402  10.26709 0.0000e+00 Sent.Completion <--- g
First.Letters     0.60838  0.072201   8.42617 0.0000e+00 First.Letters <--- g
4.Letter.Words    0.59733  0.073851   8.08843 6.6613e-16 4.Letter.Words <--- g
Suffixes          0.57179  0.071492   7.99792 1.3323e-15 Suffixes <--- g
Letter.Series     0.56689  0.074271   7.63282 2.2871e-14 Letter.Series <--- g
Pedigrees         0.66233  0.069321   9.55455 0.0000e+00 Pedigrees <--- g
Letter.Group      0.52995  0.078985   6.70955 1.9522e-11 Letter.Group <--- g
F1*Sentences      0.48787  0.085457   5.70898 1.1366e-08 Sentences <--- F1*
F1*Vocabulary     0.45232  0.090422   5.00233 5.6640e-07 Vocabulary <--- F1*
F1*Sent.Completion 0.40445 0.093402   4.33024 1.4895e-05 Sent.Completion <--- F1*
F2*First.Letters  0.61405  0.085794   7.15733 8.2268e-13 First.Letters <--- F2*
F2*4.Letter.Words 0.50581  0.084848   5.96130 2.5024e-09 4.Letter.Words <--- F2*
F2*Suffixes       0.39432  0.078289   5.03671 4.7359e-07 Suffixes <--- F2*
F3*Letter.Series  0.72730  0.159499   4.55988 5.1184e-06 Letter.Series <--- F3*
F3*Pedigrees      0.24684  0.089011   2.77317 5.5513e-03 Pedigrees <--- F3*
F3*Letter.Group   0.40915  0.122180   3.34875 8.1177e-04 Letter.Group <--- F3*
e1                0.17236  0.034113   5.05265 4.3571e-07 Sentences <--> Sentences
e2                0.16984  0.030037   5.65438 1.5641e-08 Vocabulary <--> Vocabulary
e3                0.26847  0.033188   8.08958 6.6613e-16 Sent.Completion <--> Sent.Completion
e4                0.25281  0.079472   3.18115 1.4669e-03 First.Letters <--> First.Letters
e5                0.38735  0.063194   6.12960 8.8103e-10 4.Letter.Words <--> 4.Letter.Words
e6                0.51757  0.059639   8.67838 0.0000e+00 Suffixes <--> Suffixes
e7                0.14967  0.223242   0.67044 5.0257e-01 Letter.Series <--> Letter.Series
e8                0.50039  0.059655   8.38800 0.0000e+00 Pedigrees <--> Pedigrees
e9                0.55175  0.084725   6.51223 7.4044e-11 Letter.Group <--> Letter.Group
```

14 cognitive variables from Holzinger

```
> round(Holzinger,2)

        T1   T2 T3.4   T6  T28  T29  T32  T34   T35 T36a  T13   T18  T25b   T77
T1    1.00 0.51 0.48 0.43 0.42 0.35 0.08 0.24  0.14 0.29 0.30  0.26  0.23  0.25
T2    0.51 1.00 0.66 0.50 0.40 0.43 0.15 0.25  0.08 0.37 0.55  0.53  0.44  0.43
T3.4  0.48 0.66 1.00 0.42 0.32 0.38 0.17 0.17  0.14 0.23 0.48  0.37  0.42  0.37
T6    0.43 0.50 0.42 1.00 0.44 0.53 0.06 0.37  0.21 0.39 0.35  0.35  0.31  0.28
T28   0.42 0.40 0.32 0.44 1.00 0.44 0.03 0.21  0.14 0.27 0.26  0.19  0.16  0.19
T29   0.35 0.43 0.38 0.53 0.44 1.00 0.02 0.22  0.07 0.34 0.35  0.37  0.24  0.27
T32   0.08 0.15 0.17 0.06 0.03 0.02 1.00 0.26  0.20 0.19 0.17  0.12  0.13  0.13
T34   0.24 0.25 0.17 0.37 0.21 0.22 0.26 1.00  0.33 0.44 0.20  0.16  0.05  0.04
T35   0.14 0.08 0.14 0.21 0.14 0.07 0.20 0.33  1.00 0.23 0.01 -0.01 -0.03 -0.04
T36a  0.29 0.37 0.23 0.39 0.27 0.34 0.19 0.44  0.23 1.00 0.36  0.37  0.24  0.24
T13   0.30 0.55 0.48 0.35 0.26 0.35 0.17 0.20  0.01 0.36 1.00  0.68  0.60  0.59
T18   0.26 0.53 0.37 0.35 0.19 0.37 0.12 0.16 -0.01 0.37 0.68  1.00  0.60  0.61
T25b  0.23 0.44 0.42 0.31 0.16 0.24 0.13 0.05 -0.03 0.24 0.60  0.60  1.00  0.56
T77   0.25 0.43 0.37 0.28 0.19 0.27 0.13 0.04 -0.04 0.24 0.59  0.61  0.56  1.00
>
```

How many factors

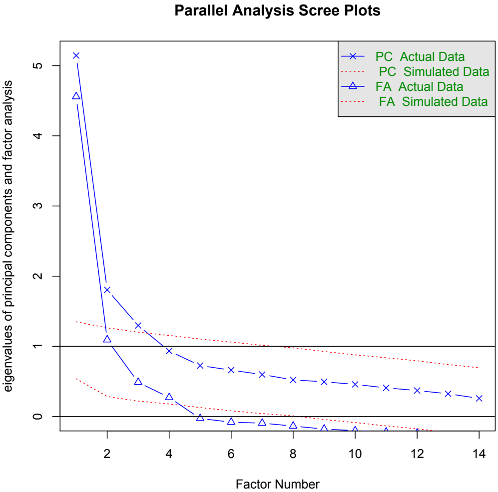
```
> vss.h <- vss(Holzinger, n.obs=355)
> vss.h

Very Simple Structure
Call: VSS(x = x, n = n, rotate = rotate, diagonal = diagonal, fm = fm,
    n.obs = n.obs, plot = plot, title = title)
VSS complexity 1 achieves a maximimum of 0.75  with  1  factors
VSS complexity 2 achieves a maximimum of 0.83  with  2  factors

The Velicer MAP criterion achieves a minimum of 0.03  with  2  factors

Velicer MAP
[1] 0.03 0.03 0.03 0.04 0.06 0.07 0.10 0.14

Very Simple Structure Complexity 1
[1] 0.75 0.59 0.56 0.51 0.49 0.48 0.44 0.44

Very Simple Structure Complexity 2
[1] 0.00 0.83 0.82 0.77 0.73 0.70 0.62 0.61
```

Chi square test

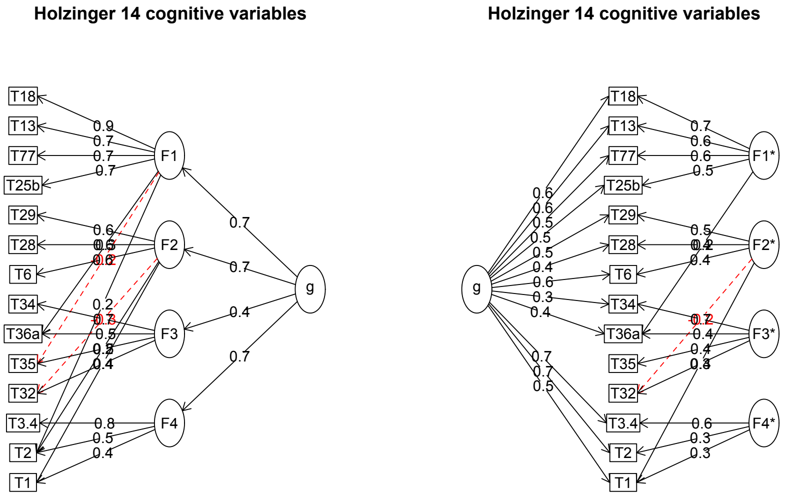
```
> print(vss.h$vss.stats[,1:4],digits=3)

  dof  chisq     prob sqresid
1  77 543.38 2.64e-71    8.84
2  64 231.86 7.26e-21    5.85
3  52 118.10 4.74e-07    4.51
4  41  41.66 4.42e-01    3.65
5  31  27.16 6.64e-01    3.45
6  22  13.58 9.16e-01    3.01
7  14   5.79 9.71e-01    2.87
8   7   2.65 9.15e-01    2.77
```

Parallel analysis of Holzinger 14 variables

> fa.parallel(Holzinger,n.obs=355) Parallel analysis suggests that the number of factors = 4 and the number of components = 3



Exploratory Omega solution for Holzinger 14 variable problem



Omega for Holzinger – exploratory and confirmatory

```
Call: omegaSem(m = Holzinger, nfactors = 4, n.obs = 355)
Omega

Alpha:                 0.85
G.6:                   0.88
Omega Hierarchical:    0.64
Omega H asymptotic:    0.71
Omega Total            0.9
```

Schmid Leiman Factor loadings greater than 0.2

	g	F1*	F2*	F3*	F4*	h2	u2	p2
T1	0.52		0.27		0.27	0.42	0.58	0.64
T2	0.69				0.34	0.66	0.34	0.72
T3.4	0.65				0.57	0.75	0.25	0.56
T6	0.57		0.42			0.54	0.46	0.59
T28	0.44		0.44			0.40	0.60	0.49
T29	0.51		0.47			0.50	0.50	0.54
T32			-0.25	0.38		0.26	0.74	0.13
T34	0.32			0.66		0.54	0.46	0.19
T35				0.44		0.25	0.75	0.10
T36a	0.43	0.20		0.44		0.45	0.55	0.41
T13	0.59	0.55				0.67	0.33	0.52
T18	0.55	0.65				0.74	0.26	0.42
T25b	0.49	0.53				0.53	0.47	0.43
T77	0.47	0.55				0.53	0.47	0.41

With eigenvalues of:

g	F1*	F2*	F3*	F4*
3.41	1.44	0.78	1.02	0.58

```
 Omega Hierarchical from a confirmatory model using se
 Omega Total  from a confirmatory model using sem =  0
With loadings of
```

	g	F1*	F2*	F3*	F4*	h2	u2
T1	0.59					0.37	0.63
T2	0.80				0.26	0.71	0.29
T3.4	0.64				0.57	0.74	0.26
T6	0.65		0.36			0.55	0.45
T28	0.50		0.34			0.36	0.64
T29	0.57		0.45			0.53	0.47
T32				0.32		0.13	0.87
T34	0.36			0.65		0.55	0.45
T35				0.44		0.22	0.78
T36a	0.51			0.38		0.41	0.59
T13	0.64	0.51				0.67	0.33
T18	0.60	0.57				0.68	0.32
T25b	0.49	0.55				0.55	0.45
T77	0.48	0.57				0.55	0.45

With eigenvalues of:

g	F1*	F2*	F3*	F4*
4.06	1.21	0.45	0.87	0.43

lavaan

- Mimics input of LISREL and MPlus
  - Easier to set up
  - Still not quite developed
  - Published version requires raw data