Psychology 454: Latent Variable Modeling
Adventures in Simulation
Testing differences between models

William Revelle

Department of Psychology
Northwestern University
Evanston, Illinois USA

October, 2016

Outline

The power of simulation

1. Simulation of structures allows us to know truth.
2. If we can recover truth, then perhaps our models work
3. If we can not recover truth, then our models do not work
4. Simulated structure
5. Simulated error

Simulated structure

1. The orthogonal case: $\boldsymbol{R} = \boldsymbol{F}\boldsymbol{F}' + \boldsymbol{U}^2$
2. The oblique case: $\boldsymbol{R} = \boldsymbol{F}\phi\boldsymbol{F}' + \boldsymbol{U}^2$
3. E = N(0,1) of dimensionality n*r
4. $\boldsymbol{R} = (\boldsymbol{F} + \boldsymbol{E})\phi(\boldsymbol{F} + \boldsymbol{E})' + \boldsymbol{U}^2$

Psych package has a number of simulation functions

1. `sim.structure` : sim.structure(fx=NULL,Phi=NULL, fy=NULL, f=NULL, n=0, uniq=NULL, raw=TRUE, items = FALSE, low=-2,high=2,d=NULL,cat=5, mu=0)
2. `sim.hierarchical`: sim.hierarchical(gload=NULL, fload=NULL, n = 0, raw = FALSE,mu = NULL)
3. `sim.congeneric`: sim.congeneric(loads = c(0.8, 0.7, 0.6, 0.5),N = NULL, err=NULL, short = TRUE, categorical=FALSE, low=-3,high=3,cuts=NULL)
4. `sim.circ` : sim.circ(nvar = 72, nsub = 500, circum = TRUE, xloading = 0.6, yloading = 0.6, gloading = 0, xbias = 0, ybias = 0, categorical = FALSE, low = -3, high = 3, truncate = FALSE, cutpoint = 0)

A congeneric model

R code

```
R <- sim.congeneric(loads=c(.9,.8,.7,.6),N=200, short=FALSE)
R
```

```
 $model (Population correlation matrix)
     V1   V2   V3   V4
V1 1.00 0.72 0.63 0.54
V2 0.72 1.00 0.56 0.48
V3 0.63 0.56 1.00 0.42
V4 0.54 0.48 0.42 1.00

$r  (Sample correlation matrix  for sample size =  200 )
     V1   V2   V3   V4
V1 1.00 0.63 0.63 0.60
V2 0.63 1.00 0.50 0.46
V3 0.63 0.50 1.00 0.43
V4 0.60 0.46 0.43 1.00
```

Exploratory Factor analysis of the congeneric model

R code

```
fm <- fa(R$model)
fm
```

```
Factor Analysis using method =  minres
Call: fa(r = R$model)
Standardized loadings (pattern matrix) based upon correlation matrix
   MR1   h2   u2 com
V1 0.9 0.81 0.19   1
V2 0.8 0.64 0.36   1
V3 0.7 0.49 0.51   1
V4 0.6 0.36 0.64   1
                MR1
SS loadings    2.30
Proportion Var 0.57
Mean item complexity =  1
Test of the hypothesis that 1 factor is sufficient.
The degrees of freedom for the null model are  6  and the objective function was  1.65
The degrees of freedom for the model are 2  and the objective function was  0
The root mean square of the residuals (RMSR) is  0
The df corrected root mean square of the residuals is  0
Fit based upon off diagonal values = 1
Measures of factor score adequacy
                                                MR1
Correlation of scores with factors             0.94
Multiple R square of scores with factors       0.88
Minimum correlation of possible factor scores  0.77
```

Exploratory Factor analysis of the congeneric data

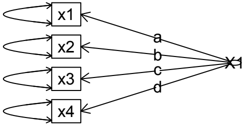
R code

```
f1 <- fa(R$observed)
f1
```

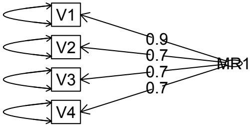
```
Factor Analysis using method =  minres
Call: fa(r = R$observed)
Standardized loadings (pattern matrix) based upon correlation matrix
    MR1   h2   u2 com
V1 0.91 0.82 0.18   1
V2 0.70 0.49 0.51   1
V3 0.69 0.48 0.52   1
V4 0.65 0.43 0.57   1
                MR1
SS loadings    2.22
Proportion Var 0.56
Mean item complexity =  1
Test of the hypothesis that 1 factor is sufficient.
The degrees of freedom for the null model are  6  and the objective function was  1.51 wi
The degrees of freedom for the model are 2  and the objective function was  0
The root mean square of the residuals (RMSR) is  0.01
The df corrected root mean square of the residuals is  0.02
The harmonic number of observations is  200 with the empirical chi square  0.29  with pro
The total number of observations was  200  with MLE Chi Square =  0.56  with prob <  0.76
Tucker Lewis Index of factoring reliability =  1.015
RMSEA index =  0  and the 90 % confidence intervals are  NA 0.095
BIC =  -10.04
Fit based upon off diagonal values = 1
Measures of factor score adequacy
                                                 MR1
Correlation of scores with factors              0.94
Multiple R square of scores with factors        0.88
```

A structure diagram of the model and of the solution

Structural model



Structural model



Getting the sem commands from `structure.diagram`

R code

```
mod <- structure.diagram(f1,errors=TRUE)
fm <- c("a","b","c","d")
mod1 <- structure.diagram(fm,errors=TRUE)
```

```
mod1
     Path        Parameter Value
[1,] "X1->x1"    "a"       NA
[2,] "X1->x2"    "b"       NA
[3,] "X1->x3"    "c"       NA
[4,] "X1->x4"    "d"       NA
[5,] "x1<->x1"   "x1e"     NA
[6,] "x2<->x2"   "x2e"     NA
[7,] "x3<->x3"   "x3e"     NA
[8,] "x4<->x4"   "x4e"     NA
[9,] "X1<->X1"   NA        "1"
```

```
mod
     Path          Parameter Value
[1,] "MR1->V1"     "F1V1"    NA
[2,] "MR1->V2"     "F1V2"    NA
[3,] "MR1->V3"     "F1V3"    NA
[4,] "MR1->V4"     "F1V4"    NA
[5,] "V1<->V1"     "x1e"     NA
[6,] "V2<->V2"     "x2e"     NA
[7,] "V3<->V3"     "x3e"     NA
[8,] "V4<->V4"     "x4e"     NA
[9,] "MR1<->MR1"   NA        "1"
```

Now do a sem on this model

R code

```
library(sem)
sem.con <- sem(mod,R$model,200)
summary(sem.con)
```

```
 Model Chisquare =  1.944223e-12   Df =  2 Pr(>Chisq) = 1
 AIC =  16
 BIC =  -10.59663

 Normalized Residuals
      Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-6.872e-07 -2.485e-07  1.293e-07  8.755e-08  5.239e-07  5.815e-07

 R-square for Endogenous Variables
  V1   V2   V3   V4
0.81 0.64 0.49 0.36

 Parameter Estimates
     Estimate Std Error  z value   Pr(>|z|)
F1V1 0.90     0.06006933 14.982688 9.528369e-51 V1 <--- MR1
F1V2 0.80     0.06278280 12.742345 3.439020e-37 V2 <--- MR1
F1V3 0.70     0.06539483 10.704210 9.725701e-27 V3 <--- MR1
F1V4 0.60     0.06808844  8.812067 1.228589e-18 V4 <--- MR1
x1e  0.19     0.04864511  3.905840 9.389880e-05 V1 <--> V1
x2e  0.36     0.05143417  6.999239 2.573568e-12 V2 <--> V2
x3e  0.51     0.05966404  8.547862 1.253891e-17 V3 <--> V3
x4e  0.64     0.06970514  9.181532 4.250011e-20 V4 <--> V4
```

sem on the actual data

R code

```
sem.con <- sem(mod,R$r,N=200)
summary(sem.con)
```

```
Model Chisquare =  0.5670591   Df =  2 Pr(>Chisq) = 0.7531209
 AIC =  16.56706
 BIC =  -10.02958

 Normalized Residuals
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
-0.270500 -0.013220  0.003423 -0.004308  0.037860  0.193400

 R-square for Endogenous Variables
    V1     V2     V3     V4
0.8240 0.4942 0.4770 0.4289

 Parameter Estimates
     Estimate  Std Error  z value   Pr(>|z|)
F1V1 0.9077513 0.06114434 14.846040 7.380138e-50 V1 <--- MR1
F1V2 0.7030014 0.06604908 10.643621 1.867291e-26 V2 <--- MR1
F1V3 0.6906628 0.06634581 10.410043 2.231199e-25 V3 <--- MR1
F1V4 0.6548883 0.06722042  9.742401 1.988009e-22 V4 <--- MR1
x1e  0.1759877 0.05380524  3.270827 1.072333e-03 V1 <--> V1
x2e  0.5057895 0.06095781  8.297370 1.064413e-16 V2 <--> V2
x3e  0.5229853 0.06201886  8.432680 3.378364e-17 V3 <--> V3
x4e  0.5711214 0.06525198  8.752553 2.085798e-18 V4 <--> V4
```

Create a data set with more subjects and then run sem

R code

```
R  <- sim.congeneric(loads=c(.9,.8,.7,.6),N=1000, short=FALSE)
sem.1000 <- sem(mod,R$r,N=1000)
summary(sem.1000)
```

```
 Model Chisquare =  0.3972647   Df =  2 Pr(>Chisq) = 0.8198513
 AIC =  16.39726
 BIC =  -13.41825

 Normalized Residuals
     Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
-0.247500 -0.023370 -0.008759 -0.011400  0.013980  0.167200

 R-square for Endogenous Variables
    V1     V2     V3     V4
0.8088 0.6417 0.4456 0.3440

 Parameter Estimates
     Estimate  Std Error  z value   Pr(>|z|)
F1V1 0.8993475 0.02710970 33.174377 2.521570e-241 V1 <--- MR1
F1V2 0.8010771 0.02822193 28.384916 3.104826e-177 V2 <--- MR1
F1V3 0.6674983 0.02968268 22.487806 5.463460e-112 V3 <--- MR1
F1V4 0.5865258 0.03061287 19.159450  8.072114e-82 V4 <--- MR1
x1e  0.1911740 0.02284898  8.366853  5.917757e-17 V1 <--> V1
x2e  0.3582755 0.02358884 15.188349  4.224229e-52 V2 <--> V2
x3e  0.5544461 0.02826857 19.613516  1.185471e-85 V3 <--> V3
x4e  0.6559873 0.03178933 20.635455  1.318996e-94 V4 <--> V4
```

Testing alternative models

1. The data were specified to be a congeneric model which allows the paths to be unequal
2. The Tau equivalent model requires the paths to be the same
3. Test the Tau equivalent model on the congeneric data

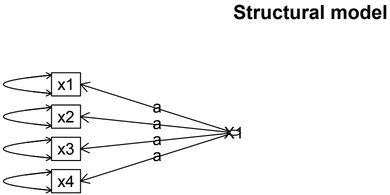
Specify the new model

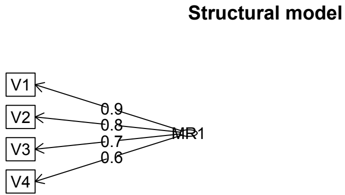
R code

```
fnew <- c("a","a","a","a")
mod.tau <- structure.diagram(fnew,errors=TRUE)
f1 <- fa(R$r)
structure.diagram(f1)
colnames(R$r) <- rownames(R$r) <- c("x1","x2","x3","x4")
mod.tau
```

```
mod.tau
      Path       Parameter Value
 [1,] "X1->x1"   "a"       NA
 [2,] "X1->x2"   "a"       NA
 [3,] "X1->x3"   "a"       NA
 [4,] "X1->x4"   "a"       NA
 [5,] "x1<->x1"  "x1e"     NA
 [6,] "x2<->x2"  "x2e"     NA
 [7,] "x3<->x3"  "x3e"     NA
 [8,] "x4<->x4"  "x4e"     NA
 [9,] "X1<->X1"  NA        "1"
attr(,"class")
```

Comparing a tau structure to the real data





The SEM analysis of the tau model with congeneric data

R code

```
sem.tau <- sem(mod.tau,R$r,N=1000)
```

```
 Model Chisquare =  101.1205   Df =  5 Pr(>Chisq) = 3.068062e-20
 AIC =  111.1205
 BIC =  66.5817

 Normalized Residuals
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-5.0310 -2.9910 -1.6410 -1.0470  0.5217  3.7070

 R-square for Endogenous Variables
    x1     x2     x3     x4
0.6758 0.6252 0.5293 0.4820

 Parameter Estimates
    Estimate  Std Error  z value  Pr(>|z|)
a    0.7695097 0.02026975 37.96346 0.000000e+00 x1 <--- X1
x1e 0.2840394 0.01864296 15.23575 2.047766e-52 x1 <--> x1
x2e 0.3550539 0.02132763 16.64760 3.149808e-62 x2 <--> x2
x3e 0.5265717 0.02838673 18.54992 8.167006e-77 x3 <--> x3
x4e 0.6364475 0.03311045 19.22196 2.424594e-82 x4 <--> x4
```

With lots of subjects., we can tell the difference

1. With lots of subjects, the model is wrong
2. What if we just had 100 subjects?
3. How about 200?

SEM with 100 subjects

R code

```
R  <- sim.congeneric(loads=c(.9,.8,.7,.6),N=100, short=FALSE)
colnames(R$r) <- rownames(R$r) <- c("x1","x2","x3","x4")
sem.100 <- sem(mod.tau,R$r,N=100)
summary(sem.100)
```

```
Model Chisquare =  8.092963   Df =  5 Pr(>Chisq) = 0.1511861
 AIC =  18.09296
 BIC =  -14.93289

 Normalized Residuals
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-1.3880 -0.7268 -0.4533 -0.2745  0.1369  1.0470

 R-square for Endogenous Variables
    x1     x2     x3     x4
0.6854 0.6367 0.5431 0.5105

 Parameter Estimates
    Estimate  Std Error  z value   Pr(>|z|)
a    0.7787127 0.06457491 12.059059 1.737560e-33 x1 <--- X1
x1e 0.2783715 0.05784667  4.812231 1.492550e-06 x1 <--> x1
x2e 0.3459893 0.06598216  5.243680 1.574053e-07 x2 <--> x2
x3e 0.5101824 0.08744696  5.834193 5.405155e-09 x3 <--> x3
x4e 0.5814753 0.09715074  5.985289 2.160059e-09 x4 <--> x4

 Iterations =  12
```

R code

```
R  <- sim.congeneric(loads=c(.9,.8,.7,.6),N=200, short=FALSE)
colnames(R$r) <- rownames(R$r) <- c("x1","x2","x3","x4")
sem.200 <- sem(mod.tau,R$r,N=200)
summary(sem.200)
```

```
 Model Chisquare =  16.6331   Df =  5 Pr(>Chisq) = 0.005250844
 AIC =  26.6331
 BIC =  -9.858483

 Normalized Residuals
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-1.9600 -1.6120 -0.1834 -0.3544  0.4208  1.4400

 R-square for Endogenous Variables
    x1     x2     x3     x4
0.6602 0.6526 0.5823 0.4873

 Parameter Estimates
    Estimate  Std Error  z value   Pr(>|z|)
a   0.7787937 0.04553742 17.102281 1.427077e-65 x1 <--- X1
x1e 0.3122247 0.04356330  7.167149 7.657588e-13 x1 <--> x1
x2e 0.3228945 0.04447735  7.259751 3.878022e-13 x2 <--> x2
x3e 0.4350640 0.05457012  7.972567 1.554117e-15 x3 <--> x3
x4e 0.6380082 0.07398931  8.622978 6.523512e-18 x4 <--> x4
```

Congeneric for 9 items

R code

```
f <- c(.9,.8,.7,.8,.7,.6,.7,.6,.5)
R <- sim.congeneric(loads=f)
f1 <- fa(R)
f1
```

```
Factor Analysis using method =  minres
Call: fa(r = R)
Standardized loadings (pattern matrix) based upon correlation matrix
   MR1   h2   u2 com
V1 0.9 0.81 0.19   1
V2 0.8 0.64 0.36   1
V3 0.7 0.49 0.51   1
V4 0.8 0.64 0.36   1
V5 0.7 0.49 0.51   1
V6 0.6 0.36 0.64   1
V7 0.7 0.49 0.51   1
V8 0.6 0.36 0.64   1
V9 0.5 0.25 0.75   1

                MR1
SS loadings    4.53
Proportion Var 0.50

Mean item complexity =  1
Test of the hypothesis that 1 factor is sufficient.

The degrees of freedom for the null model are  36  and the objective function was  4.33
The degrees of freedom for the model are 27  and the objective function was  0
```

Add some error to the 9 variable congeneric model

R code

```
f <- c(.9,.8,.7,.8,.7,.6,.7,.6,.5)
R <- sim.congeneric(loads=f, N=200,short=FALSE)
f1 <- fa(R$r)
f1
```

```
Factor Analysis using method =  minres
Call: fa(r = R$r)
Standardized loadings (pattern matrix) based upon correlation matrix
    MR1   h2   u2 com
V1 0.92 0.85 0.15   1
V2 0.77 0.59 0.41   1
V3 0.66 0.43 0.57   1
V4 0.78 0.61 0.39   1
V5 0.68 0.46 0.54   1
V6 0.60 0.35 0.65   1
V7 0.69 0.48 0.52   1
V8 0.63 0.40 0.60   1
V9 0.51 0.26 0.74   1
                MR1
SS loadings    4.44
Proportion Var 0.49
Mean item complexity =  1
Test of the hypothesis that 1 factor is sufficient.
The degrees of freedom for the null model are  36  and the objective function was  4.29
The degrees of freedom for the model are 27  and the objective function was  0.09
The root mean square of the residuals (RMSR) is  0.03
The df corrected root mean square of the residuals is  0.03
Fit based upon off diagonal values = 1
```

Try a SEM/CFA

R code

```
mod.9 <- structure.diagram(f1,errors=TRUE)
mod.9
sem.9 <- sem(mod.9,R$r,N=200)
summary(sem.9)
```

```
      Path         Parameter Value
 [1,] "MR1->V1"    "F1V1"    NA
 [2,] "MR1->V2"    "F1V2"    NA
 [3,] "MR1->V3"    "F1V3"    NA
 [4,] "MR1->V4"    "F1V4"    NA
 [5,] "MR1->V5"    "F1V5"    NA
 [6,] "MR1->V6"    "F1V6"    NA
 [7,] "MR1->V7"    "F1V7"    NA
 [8,] "MR1->V8"    "F1V8"    NA
 [9,] "MR1->V9"    "F1V9"    NA
[10,] "V1<->V1"    "x1e"     NA
[11,] "V2<->V2"    "x2e"     NA
[12,] "V3<->V3"    "x3e"     NA
[13,] "V4<->V4"    "x4e"     NA
[14,] "V5<->V5"    "x5e"     NA
[15,] "V6<->V6"    "x6e"     NA
[16,] "V7<->V7"    "x7e"     NA
[17,] "V8<->V8"    "x8e"     NA
[18,] "V9<->V9"    "x9e"     NA
[19,] "MR1<->MR1" NA         "1"
```

The CFA results

R code

```
summary(sem.9)
```

```
 Model Chisquare =  16.99844   Df =  27 Pr(>Chisq) = 0.9311608
 AIC =  52.99844
 BIC =  -126.0561

 Normalized Residuals
      Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-0.7174000 -0.1553000  0.0000005  0.0049640  0.2014000  0.7404000

 R-square for Endogenous Variables
    V1     V2     V3     V4     V5     V6     V7     V8     V9
0.8543 0.5892 0.4297 0.6054 0.4644 0.3546 0.4822 0.3960 0.2636

 Parameter Estimates
     Estimate  Std Error  z value   Pr(>|z|)
F1V1 0.9242876 0.05506791 16.784504 3.168638e-63 V1 <--- MR1
F1V2 0.7676013 0.06119648 12.543227 4.329960e-36 V2 <--- MR1
F1V3 0.6555021 0.06475709 10.122475 4.391675e-24 V3 <--- MR1
F1V4 0.7780761 0.06082697 12.791629 1.826050e-37 V4 <--- MR1
F1V5 0.6814464 0.06399641 10.648196 1.777763e-26 V5 <--- MR1
F1V6 0.5954412 0.06637832  8.970418 2.953916e-19 V6 <--- MR1
F1V7 0.6944168 0.06360204 10.918151 9.439759e-28 V7 <--- MR1
F1V8 0.6292949 0.06548811  9.609300 7.304373e-22 V8 <--- MR1
F1V9 0.5134170 0.06829351  7.517801 5.570523e-14 V9 <--- MR1
x1e  0.1456921 0.02718597  5.359092 8.364108e-08 V1 <--> V1
x2e  0.4107882 0.04656243  8.822309 1.121229e-18 V2 <--> V2
x3e  0.5703168 0.06078765  9.382116 6.466150e-21 V3 <--> V3
x4e  0.3945974 0.04515869  8.738018 2.372356e-18 V4 <--> V4
x5e  0.5356307 0.05765381  9.290465 1.536164e-20 V5 <--> V5
```

Now, simulate a more complicated (hierarchical) model

1. 9 variables (e.g., 9 measures of Neuroticism or 9 measures of ability)
2. 3 group factors (e.g., anxiety, depression, aggression)
3. 1 higher order factor (e.g. Neuroticism)
4. Simulate using (sim.hierarchical) or could use `sim.structural`

Create a 9 variable hierarchical model

R code

```
nine <- sim.hierarchical()
 lowerMat(nine)
```

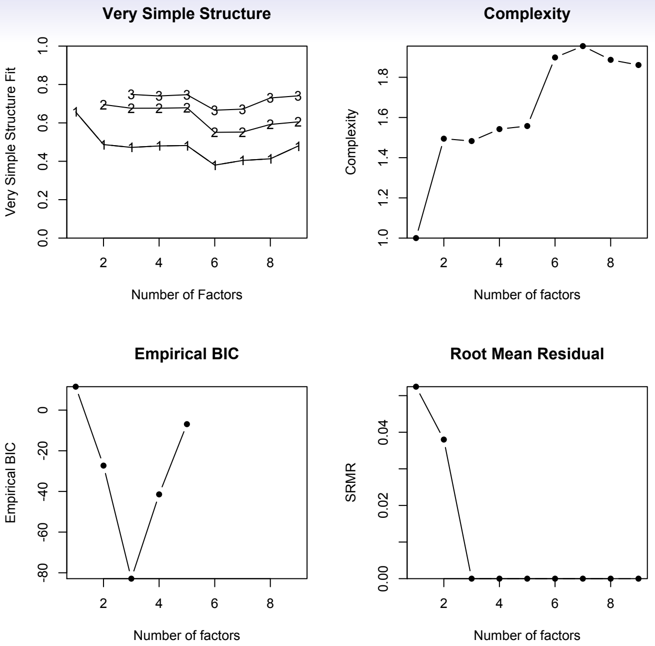
```
   V1   V2   V3   V4   V5   V6   V7   V8   V9
V1 1.00
V2 0.56 1.00
V3 0.48 0.42 1.00
V4 0.40 0.35 0.30 1.00
V5 0.35 0.30 0.26 0.42 1.00
V6 0.29 0.25 0.22 0.35 0.30 1.00
V7 0.30 0.26 0.23 0.24 0.20 0.17 1.00
V8 0.25 0.22 0.19 0.20 0.17 0.14 0.30 1.00
V9 0.20 0.18 0.15 0.16 0.13 0.11 0.24 0.20 1.00
```

How many factors are really in the data?

R code

```
nfactors(nine)
```

```
Number of factors
Call: vss(x = x, n = n, rotate = rotate, diagonal = diagonal, fm = fm,
    n.obs = n.obs, plot = FALSE, title = title, use = use, cor = cor)
VSS complexity 1 achieves a maximimum of 0.66  with  1  factors
VSS complexity 2 achieves a maximimum of 0.7  with  2  factors
The Velicer MAP achieves a minimum of 0.02  with  1  factors
Empirical BIC achieves a minimum of  -82.89  with  3  factors
Sample Size adjusted BIC achieves a minimum of  -44.78  with  3  factors

Statistics by number of factors
  vss1 vss2   map dof   chisq    prob sqresid  fit RMSEA   BIC SABIC complex  eChisq    S
1 0.66 0.00 0.023  27 1.6e+02 4.2e-21     5.1 0.66 0.071 -25.1  60.7     1.0 2.0e+02 5.2e
2 0.49 0.70 0.042  19 6.9e+01 1.2e-07     4.5 0.70 0.052 -61.9  -1.6     1.5 1.0e+02 3.8e
3 0.47 0.68 0.064  12 2.8e-07 1.0e+00     3.7 0.75 0.000 -82.9 -44.8     1.5 2.8e-07 2.0e
4 0.48 0.68 0.127   6 1.8e-07 1.0e+00     3.6 0.76 0.000 -41.4 -22.4     1.5 1.8e-07 1.6e
5 0.48 0.68 0.189   1 2.4e-09 1.0e+00     3.5 0.76 0.000  -6.9  -3.7     1.6 3.2e-09 2.1e
6 0.38 0.55 0.330  -3 3.0e-09      NA     3.3 0.78    NA    NA    NA     1.9 3.9e-09 2.3e
7 0.40 0.55 0.457  -6 7.5e-07      NA     3.2 0.78    NA    NA    NA     2.0 9.8e-07 3.7e
8 0.41 0.59 1.000  -8 0.0e+00      NA     3.1 0.79    NA    NA    NA     1.9 2.0e-17 1.7e
9 0.48 0.61    NA  -9 0.0e+00      NA     3.1 0.79    NA    NA    NA     1.9 1.5e-26 4.5e
```


Very Simple Structure
Very Simple Structure Fit
Number of Factors
Complexity
Complexity
Number of factors
Empirical BIC
Empirical BIC
Number of factors
Root Mean Residual
SRMR
Number of factors

What about 1 factor?

R code

```
f1 <- fa(nine)
f1
```

```
Factor Analysis using method =  minres
Call: fa(r = nine)
Standardized loadings (pattern matrix) based upon correlation matrix
    MR1    h2   u2 com
V1 0.75 0.560 0.44   1
V2 0.67 0.452 0.55   1
V3 0.58 0.339 0.66   1
V4 0.58 0.332 0.67   1
V5 0.51 0.256 0.74   1
V6 0.43 0.183 0.82   1
V7 0.43 0.182 0.82   1
V8 0.36 0.131 0.87   1
V9 0.29 0.086 0.91   1

                MR1
SS loadings    2.52
Proportion Var 0.28

Mean item complexity =  1
Test of the hypothesis that 1 factor is sufficient.

The degrees of freedom for the null model are  36  and the objective function was  1.71
The degrees of freedom for the model are 27  and the objective function was  0.16

The root mean square of the residuals (RMSR) is  0.05
The df corrected root mean square of the residuals is  0.06
```

Try a 1 factor CFA– get the model

R code

```
mod1 <- structure.diagram(f1,cut=.1,errors=TRUE)  #must specify er
 mod1
```

```
mod1
      Path         Parameter Value
 [1,] "MR1->V1"    "F1V1"    NA
 [2,] "MR1->V2"    "F1V2"    NA
 [3,] "MR1->V3"    "F1V3"    NA
 [4,] "MR1->V4"    "F1V4"    NA
 [5,] "MR1->V5"    "F1V5"    NA
 [6,] "MR1->V6"    "F1V6"    NA
 [7,] "MR1->V7"    "F1V7"    NA
 [8,] "MR1->V8"    "F1V8"    NA
 [9,] "MR1->V9"    "F1V9"    NA
[10,] "V1<->V1"    "x1e"     NA
[11,] "V2<->V2"    "x2e"     NA
[12,] "V3<->V3"    "x3e"     NA
[13,] "V4<->V4"    "x4e"     NA
[14,] "V5<->V5"    "x5e"     NA
[15,] "V6<->V6"    "x6e"     NA
[16,] "V7<->V7"    "x7e"     NA
[17,] "V8<->V8"    "x8e"     NA
[18,] "V9<->V9"    "x9e"     NA
[19,] "MR1<->MR1" NA         "1"
```

Test the 1 factor CFA– it works! Even though not the the model

R code

```
sem1 <- sem(mod1,nine,N=200)
summary(sem1)
```

```
Model Chisquare =  32.2424   Df =  27 Pr(>Chisq) = 0.2232256
 AIC =  68.2424
 BIC =  -110.8122

 Normalized Residuals
     Min.    1st Qu.     Median       Mean    3rd Qu.      Max.
-0.5161000 -0.3010000 -0.1678000  0.0676700  0.0000032  2.0650000

 R-square for Endogenous Variables
    V1     V2     V3     V4     V5     V6     V7     V8     V9
0.5731 0.4596 0.3424 0.3298 0.2531 0.1805 0.1790 0.1290 0.0847

 Parameter Estimates
     Estimate  Std Error  z value   Pr(>|z|)
F1V1 0.7570531 0.06741243 11.230170 2.899217e-29 V1 <--- MR1
F1V2 0.6779741 0.06940423  9.768483 1.537372e-22 V2 <--- MR1
F1V3 0.5851384 0.07171144  8.159624 3.360702e-16 V3 <--- MR1
F1V4 0.5742815 0.07196977  7.979482 1.469481e-15 V4 <--- MR1
F1V5 0.5030470 0.07356384  6.838237 8.017390e-12 V5 <--- MR1
F1V6 0.4248877 0.07508254  5.658941 1.523098e-08 V6 <--- MR1
F1V7 0.4230254 0.07511561  5.631657 1.784866e-08 V7 <--- MR1
F1V8 0.3591109 0.07616124  4.715140 2.415450e-06 V8 <--- MR1
F1V9 0.2910727 0.07708303  3.776093 1.593074e-04 V9 <--- MR1
x1e  0.4268707 0.06348726  6.723723 1.771396e-11 V1 <--> V1
x2e  0.5403512 0.06837469  7.902796 2.727161e-15 V2 <--> V2
x3e  0.6576134 0.07540062  8.721591 2.743188e-18 V3 <--> V3
x4e  0.6702004 0.07623152  8.791644 1.473873e-18 V4 <--> V4
```

An exploratory solution with correlated factors

R code

```
f3 <- fa(nine,3)
 f3
```

```
Factor Analysis using method =  minres
Call: fa(r = nine, nfactors = 3)
Standardized loadings (pattern matrix) based upon correlation matrix
   MR1 MR2 MR3   h2   u2 com
V1 0.8 0.0 0.0 0.64 0.36   1
V2 0.7 0.0 0.0 0.49 0.51   1
V3 0.6 0.0 0.0 0.36 0.64   1
V4 0.0 0.7 0.0 0.49 0.51   1
V5 0.0 0.6 0.0 0.36 0.64   1
V6 0.0 0.5 0.0 0.25 0.75   1
V7 0.0 0.0 0.6 0.36 0.64   1
V8 0.0 0.0 0.5 0.25 0.75   1
V9 0.0 0.0 0.4 0.16 0.84   1

                       MR1  MR2  MR3
SS loadings           1.49 1.10 0.77
Proportion Var        0.17 0.12 0.09
Cumulative Var        0.17 0.29 0.37
Proportion Explained  0.44 0.33 0.23
Cumulative Proportion 0.44 0.77 1.00

 With factor correlations of
     MR1  MR2  MR3
MR1 1.00 0.72 0.63
MR2 0.72 1.00 0.56
MR3 0.63 0.56 1.00
```

What about a bi-factor solution?

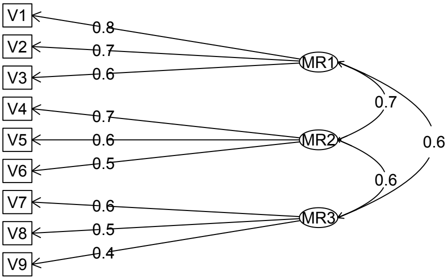
R code

```
f3.bi <- fa(nine,3,rotate="bifactor")
 f3.bi
```

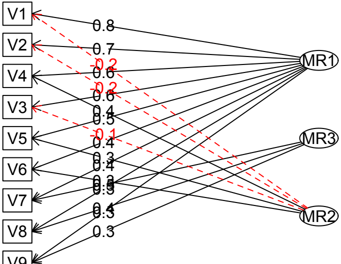
```
Factor Analysis using method =  minres
Call: fa(r = nine, nfactors = 3, rotate = "bifactor")
Standardized loadings (pattern matrix) based upon correlation matrix
    MR1  MR3   MR2   h2   u2 com
V1 0.78 0.00 -0.19 0.64 0.36 1.1
V2 0.68 0.00 -0.17 0.49 0.51 1.1
V3 0.58 0.00 -0.14 0.36 0.64 1.1
V4 0.60 0.00  0.35 0.49 0.51 1.6
V5 0.52 0.00  0.30 0.36 0.64 1.6
V6 0.43 0.00  0.25 0.25 0.75 1.6
V7 0.39 0.46  0.00 0.36 0.64 2.0
V8 0.32 0.38  0.00 0.25 0.75 2.0
V9 0.26 0.30  0.00 0.16 0.84 2.0

                       MR1  MR3  MR2
SS loadings           2.55 0.45 0.36
Proportion Var        0.28 0.05 0.04
Cumulative Var        0.28 0.33 0.37
Proportion Explained  0.76 0.13 0.11
Cumulative Proportion 0.76 0.89 1.00

Mean item complexity =  1.6
Test of the hypothesis that 3 factors are sufficient.
```

Comparing an oblique to a bifactor solution

3 oblique factors



A bifactor solution



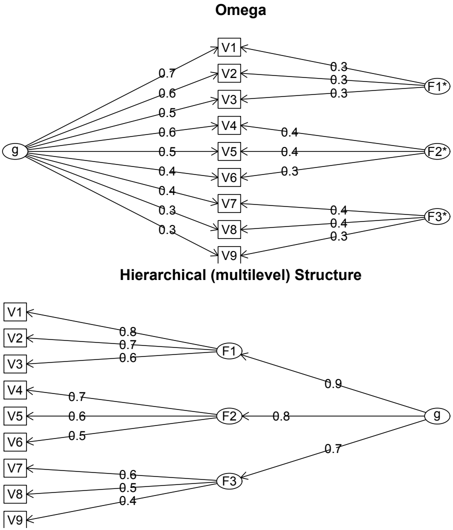
Try an explicitly hierarchical solution

R code

```
om <- omega(nine)  #by default, draws the Schmid-Leiman solution
omega.diagram(om,sl=FALSE) #draws the hierarchical solution
```

```
om
Omega
Call: omega(m = nine)
Alpha:                 0.76
G.6:                   0.76
Omega Hierarchical:    0.69
Omega H asymptotic:    0.86
Omega Total            0.8

Schmid Leiman Factor loadings greater than  0.2
      g  F1*  F2*  F3*  h2   u2   p2
V1 0.72 0.35            0.64 0.36 0.81
V2 0.63 0.31            0.49 0.51 0.81
V3 0.54 0.26            0.36 0.64 0.81
V4 0.56      0.42       0.49 0.51 0.64
V5 0.48      0.36       0.36 0.64 0.64
V6 0.40      0.30       0.25 0.75 0.64
V7 0.42           0.43 0.36 0.64 0.49
V8 0.35           0.36 0.25 0.75 0.49
V9 0.28           0.29 0.16 0.84 0.49
```

The omega function draws both the Schmid-Leiman and the hierarchical soution



The model for the higher order solution

R code

```
om$model
```

```
      Path          Parameter Initial Value
 [1,] "g->V1"       "V1"       NA
 [2,] "g->V2"       "V2"       NA
 [3,] "g->V3"       "V3"       NA
 [4,] "g->V4"       "V4"       NA
 [5,] "g->V5"       "V5"       NA
 [6,] "g->V6"       "V6"       NA
 [7,] "g->V7"       "V7"       NA
 [8,] "g->V8"       "V8"       NA
 [9,] "g->V9"       "V9"       NA
[10,] "F1*->V1"     "F1*V1"    NA
[11,] "F1*->V2"     "F1*V2"    NA
[12,] "F1*->V3"     "F1*V3"    NA
[13,] "F2*->V4"     "F2*V4"    NA
[14,] "F2*->V5"     "F2*V5"    NA
[15,] "F2*->V6"     "F2*V6"    NA
[16,] "F3*->V7"     "F3*V7"    NA
[17,] "F3*->V8"     "F3*V8"    NA
[18,] "F3*->V9"     "F3*V9"    NA
[19,] "V1<->V1"     "e1"       NA
[20,] "V2<->V2"     "e2"       NA
[21,] "V3<->V3"     "e3"       NA
[22,] "V4<->V4"     "e4"       NA
[23,] "V5<->V5"     "e5"       NA
[24,] "V6<->V6"     "e6"       NA
[25,] "V7<->V7"     "e7"       NA
[26,] "V8<->V8"     "e8"       NA
[27,] "V9<->V9"     "e9"       NA
```

The CFA

R code

```
sem.om <- sem(om$model,nine,N=100)
```

```
 sem.om
 Model Chisquare =  1.85004e-10   Df =  18 Pr(>Chisq) = 1
 AIC =  54
 BIC =  -82.89306

 Normalized Residuals
      Min.    1st Qu.     Median       Mean    3rd Qu.       Max.
-4.925e-06 -1.696e-06 -8.395e-07 -4.049e-07  7.230e-07  3.897e-06

 R-square for Endogenous Variables
  V1   V2   V3   V4   V5   V6   V7   V8   V9
0.64 0.49 0.36 0.49 0.36 0.25 0.36 0.25 0.16

 Parameter Estimates
      Estimate  Std Error z value  Pr(>|z|)
V1    0.7200003 0.1337326 5.383881 7.289689e-08 V1 <--- g
V2    0.6299999 0.1352950 4.656490 3.216464e-06 V2 <--- g
V3    0.5400003 0.1366347 3.952147 7.745323e-05 V3 <--- g
V4    0.5599999 0.1228772 4.557394 5.179212e-06 V4 <--- g
V5    0.4800000 0.1229732 3.903290 9.489380e-05 V5 <--- g
V6    0.4000001 0.1230543 3.250597 1.151629e-03 V6 <--- g
V7    0.4199999 0.1163253 3.610564 3.055320e-04 V7 <--- g
V8    0.3500004 0.1174465 2.980083 2.881701e-03 V8 <--- g
V9    0.2799994 0.1183560 2.365738 1.799416e-02 V9 <--- g
F1*V1 0.3487112 0.2789614 1.250034 2.112873e-01 V1 <--- F1*
F1*V2 0.3051237 0.2674266 1.140962 2.538856e-01 V2 <--- F1*
F1*V3 0.2615333 0.2570111 1.017595 3.088703e-01 V3 <--- F1*
```