Psychology 454: Latent Variable Modeling

Week 3: Adventures in simulation

Department of Psychology
Northwestern University
Evanston, Illinois USA

January, 2011

Outline

Adventures in simulation

- To understand particular structures, it is useful to know "truth".
- Can generate a particular structure and see how various algorithms recover it.
    - Exploratory factoring
    - Confirmatory factoring
    - For more complicated structures, structural equation modeling
- For all models, we can evaluate the various goodness of fit statistics
- What happens if we create data that intentionally violate the assumptions

The basic model

- Factor model: $\mathbf{R} = \mathbf{F}\phi\mathbf{F'} + \mathbf{U}^2$
- Covariance model: $\mathbf{C} = \mathbf{P}\phi\mathbf{P'} + \mathbf{U}^2$
    - Model is Pattern * Phi * t(Pattern) + error
- Let $\mathbf{S}$ represent the observed covariance matrix,
- Let Σ be the modeled matrix with $\Sigma = \mathbf{FF' + U^2}$, then minimize

$$E = \frac{1}{2} tr(\mathbf{S} - \Sigma)^2 \tag{1}$$

Varieties of minimization

$$E = \frac{1}{2} tr((\mathbf{S} - \Sigma)\mathbf{S}^{-1})^2 = \frac{1}{2} tr(\mathbf{I} - \Sigma \mathbf{S}^{\mathbf{-1}})^2. \tag{2}$$

This is known as *generalized least squares* (*GLS*) or *weighted least squares* (*WLS*).
Similarly, if the residuals are weighted by the inverse of the model, Σ, minimizing

$$E = \frac{1}{2} tr((\mathbf{S} - \Sigma)\Sigma^{-1})^2 = \frac{1}{2} tr(\mathbf{S}\Sigma^{-1} - I)^2 \tag{3}$$

will result in a model that maximizes the likelihood of the data.
This procedure, *maximum likelihood estimation* (*MLE*) is also seen as finding the minimum of

$$E = \frac{1}{2} \left(tr(\Sigma^{-1}\mathbf{S}) - ln \left|\Sigma^{-1}\mathbf{S}\right| - p \right) \tag{4}$$

Simulate the model using sim function

```
> fx <- matrix(c(.9,.8,.7,rep(0 ,9),.9,.8,.7,rep(0,9),.9,.8,.7),9,3)
> fx
      [,1] [,2] [,3]
 [1,]  0.9  0.0  0.0
 [2,]  0.8  0.0  0.0
 [3,]  0.7  0.0  0.0
 [4,]  0.0  0.9  0.0
 [5,]  0.0  0.8  0.0
 [6,]  0.0  0.7  0.0
 [7,]  0.0  0.0  0.9
 [8,]  0.0  0.0  0.8
 [9,]  0.0  0.0  0.7
> Phi <- diag(3)
> R <- sim(fx,Phi)
> R
```

```
Call: sim(fx = fx, Phi = Phi)

 $model (Population correlation matrix)
     V1   V2   V3   V4   V5   V6   V7   V8   V9
V1 1.00 0.72 0.63 0.00 0.00 0.00 0.00 0.00 0.00
V2 0.72 1.00 0.56 0.00 0.00 0.00 0.00 0.00 0.00
V3 0.63 0.56 1.00 0.00 0.00 0.00 0.00 0.00 0.00
V4 0.00 0.00 0.00 1.00 0.72 0.63 0.00 0.00 0.00
V5 0.00 0.00 0.00 0.72 1.00 0.56 0.00 0.00 0.00
V6 0.00 0.00 0.00 0.63 0.56 1.00 0.00 0.00 0.00
V7 0.00 0.00 0.00 0.00 0.00 0.00 1.00 0.72 0.63
V8 0.00 0.00 0.00 0.00 0.00 0.00 0.72 1.00 0.56
V9 0.00 0.00 0.00 0.00 0.00 0.00 0.63 0.56 1.00

$reliability (population reliability)
[1] 0.81 0.64 0.49 0.81 0.64 0.49 0.81 0.64 0.49
```

Simulate 400 subjects

```
> R <- sim(fx,Phi,n=400)
> R
Call: sim(fx = fx, Phi = Phi, n = 400)

...
$r  (Sample correlation matrix  for sample size =  400 )
        V1      V2     V3     V4      V5     V6     V7     V8      V9
V1  1.0000  0.6839  0.609 -0.020 -0.0328  0.024  0.093  0.074  0.0061
V2  0.6839  1.0000  0.532 -0.025 -0.0093 -0.025  0.143  0.111  0.0726
V3  0.6088  0.5316  1.000 -0.040 -0.0362  0.062  0.073  0.030  0.0070
V4 -0.0196 -0.0253 -0.040  1.000  0.7379  0.646  0.066  0.093  0.0113
V5 -0.0328 -0.0093 -0.036  0.738  1.0000  0.581  0.058  0.096  0.0840
V6  0.0245 -0.0247  0.062  0.646  0.5813  1.000 -0.038 -0.043 -0.0636
V7  0.0927  0.1434  0.073  0.066  0.0580 -0.038  1.000  0.721  0.6193
V8  0.0738  0.1111  0.030  0.093  0.0962 -0.043  0.721  1.000  0.5460
V9  0.0061  0.0726  0.007  0.011  0.0840 -0.064  0.619  0.546  1.0000
>
```

factor it nf=1

```
> summary(fa(R$observed))


Factor analysis with Call: fa(r = R$observed)

Test of the hypothesis that 1 factor is sufficient.
The degrees of freedom for the model is 27  and the
                  objective function was  2.9
The number of observations was  400  with Chi Square =
                              1144.98  with prob <  4.2e-224

The root mean square of the residuals is  0.19
The df corrected root mean square of the residuals is  0.31

Tucker Lewis Index of factoring reliability =  -0.002
RMSEA index =  0.324  and the 90 % confidence intervals are  0.323 0.325
BIC =  983.21
```

Try two factors

```
> f2 <- fa(R$observed,2)
> summary(f2)

Factor analysis with Call: fa(r = R$observed, nfactors = 2)

Test of the hypothesis that 2 factors are sufficient.
The degrees of freedom for the model is 19  and the
                             objective function was  1.18
The number of observations was  400  with Chi Square =
                                   465.42  with prob <  9.8e-87

The root mean square of the residuals is  0.12
The df corrected root mean square of the residuals is  0.24

Tucker Lewis Index of factoring reliability =  0.431
RMSEA index =  0.244  and the 90 % confidence intervals are  0.243 0.246
BIC =  351.59
>
```

How about 3

```
> f3 <- fa(R$observed,3)
> summary(f3)
```

```
Factor analysis with Call: fa(r = R$observed, nfactors = 3)

Test of the hypothesis that 3 factors are sufficient.
The degrees of freedom for the model is 12  and
          the objective function was  0.05
The number of observations was  400  with Chi Square =
               18.96  with prob <  0.089

The root mean square of the residuals is  0.01
The df corrected root mean square of the residuals is  0.03

Tucker Lewis Index of factoring reliability =  0.986
RMSEA index =  0.039  and the 90 % confidence intervals are  0.039 0.043
BIC =  -52.93
```

That looks better, lets look at the structure

```
> f3
Factor Analysis using method =  minres
Call: fa(r = R$observed, nfactors = 3)
Standardized loadings based upon correlation matrix
     MR1   MR2   MR3   h2   u2
V1  0.01 -0.02  0.89 0.78 0.22
V2 -0.01  0.06  0.77 0.60 0.40
V3 -0.01 -0.02  0.69 0.48 0.52
V4  0.90  0.02 -0.01 0.82 0.18
V5  0.81  0.04 -0.02 0.66 0.34
V6  0.72 -0.10  0.05 0.53 0.47
V7 -0.01  0.89  0.03 0.80 0.20
V8  0.03  0.80  0.00 0.65 0.35
V9 -0.03  0.70 -0.05 0.48 0.52

                MR1  MR2  MR3                         With factor correlations of

SS loadings    2.00 1.94 1.86                            MR1  MR2   MR3
Proportion Var 0.22 0.22 0.21                        MR1  1.00 0.06 -0.02
Cumulative Var 0.22 0.44 0.64                        MR2  0.06 1.00  0.11
                                                     MR3 -0.02 0.11  1.00
```

Try a confirmatory – using lavaan

```
libary(lavaan)
my.data <- data.frame(R$observed)  #weird but necessary
model <- 'F1 =~ V1 + V2 + V3
                F2 =~ V4 + V5 + V6
                F3 =~ V7 + V8 + V9'
fit <- cfa(model,data = my.data)
> fit
```

```
Lavaan (0.4-5) converged normally after 25 iterations

  Number of observations                           400

  Estimator                                         ML
  Minimum Function Chi-square                   37.832
  Degrees of freedom                                24
  P-value                                        0.036
```

Compare this to the exploratory – more df, not quite as good a fit.

Examine the structure

```
> summary(fit)

                   Estimate  Std.err  Z-value  P(>|z|)
Latent variables:
  F1 =~
    V1                1.000
    V2                0.905    0.064   14.226    0.000
    V3                0.753    0.057   13.180    0.000
  F2 =~
    V4                1.000
    V5                0.920    0.055   16.762    0.000
    V6                0.736    0.049   14.939    0.000
  F3 =~
    V7                1.000
    V8                0.831    0.054   15.414    0.000
    V9                0.734    0.054   13.690    0.000

Covariances:
  F1 ~~
    F2               -0.021    0.045   -0.463    0.643
    F3                0.093    0.044    2.098    0.036
  F2 ~~
    F3                0.056    0.047    1.205    0.228

Variances:
    V1                0.208    0.042    4.912    0.000
    V2                0.388    0.043    9.099    0.000
    ...
    V9                0.495    0.041   11.991    0.000
    F1                0.721    0.075    9.562    0.000
    F2                0.836    0.080   10.434    0.000
```

Different ways of fixing variances

```
fit <- cfa(model,data =my.data,std.lv=TRUE)  #set latents to have variance 1
                   Estimate  Std.err  Z-value  P(>|z|)
Latent variables:
  F1 =~
    V1                0.849    0.044   19.124    0.000
    V2                0.768    0.046   16.539    0.000
    V3                0.640    0.044   14.450    0.000
  F2 =~
    V4                0.915    0.044   20.868    0.000
    V5                0.841    0.046   18.187    0.000
    V6                0.673    0.044   15.438    0.000
  F3 =~
    V7                0.896    0.044   20.300    0.000
    V8                0.744    0.043   17.336    0.000
    V9                0.657    0.045   14.504    0.000

Covariances:
  F1 ~~
    F2               -0.027    0.057   -0.464    0.643
    F3                0.122    0.057    2.142    0.032
  F2 ~~
    F3                0.069    0.057    1.213    0.225

Variances:
    V1                0.208    0.042    4.912    0.000
    V2                0.388    0.043    9.099    0.000
   ...
    V8                0.320    0.036    8.898    0.000
    V9                0.495    0.041   11.991    0.000
    F1                1.000
    F2                1.000
    F3                1.000
```

Goodness of fit and sample size

- If the model actually fits, increasing sample size does not harm goodness of fit
  - This is if the model is in fact a perfect model of the data
  - If the model does not fit, then χ^2 will get worse
- Other goodness of fit statistics are not as sensitive to sample size.
  - But this is relevant if the model is in fact incorrect

Vary the sample size

```
> set.seed(42)
> R <- sim(fx,Phi,n=100)
> summary(fa(R$observed,3))
```

```
Factor analysis with Call: fa(r = R$observed, nfactors = 3)

Test of the hypothesis that 3 factors are sufficient.
The degrees of freedom for the model is 12  and the
      objective function was  0.11
The number of observations was  100  with
      Chi Square =  10.42  with prob <  0.58

The root mean square of the residuals is  0.02
The df corrected root mean square of the residuals is  0.04

Tucker Lewis Index of factoring reliability =  1.013
RMSEA index =  0  and the 90 % confidence intervals are  0 0.034
BIC =  -44.84
```

N =400

```
> set.seed(42)
> R <- sim(fx,Phi,n=400)
> summary(fa(R$observed,3))
```

```
Factor analysis with Call: fa(r = R$observed, nfactors = 3)

Test of the hypothesis that 3 factors are sufficient.
The degrees of freedom for the model is 12  and the
       objective function was  0.07
The number of observations was  400  with
        Chi Square =  26.18  with prob <  0.01

The root mean square of the residuals is  0.01
The df corrected root mean square of the residuals is  0.03

Tucker Lewis Index of factoring reliability =  0.971
RMSEA index =  0.055  and the 90 % confidence intervals are  0.055 0.059
BIC =  -45.72
```

N = 800

```
> set.seed(42)
> R <- sim(fx,Phi,n=800)
> summary(fa(R$observed,3))
```

```
Factor analysis with Call: fa(r = R$observed, nfactors = 3)

Test of the hypothesis that 3 factors are sufficient.
The degrees of freedom for the model is 12  and the
       objective function was  0.03
The number of observations was  800  with
       Chi Square =  22.4  with prob <  0.033

The root mean square of the residuals is  0.01
The df corrected root mean square of the residuals is  0.02

Tucker Lewis Index of factoring reliability =  0.99
RMSEA index =  0.033  and the 90 % confidence intervals are  0.033 0.036
BIC =  -57.82
```

N =1600

```
> set.seed(42)
> R <- sim(fx,Phi,n=1600)
> summary(fa(R$observed,3))
```

```
Factor analysis with Call: fa(r = R$observed, nfactors = 3)

Test of the hypothesis that 3 factors are sufficient.
The degrees of freedom for the model is 12  and the
       objective function was  0.01
The number of observations was  1600  with
       Chi Square =  21.65  with prob <  0.042

The root mean square of the residuals is  0.01
The df corrected root mean square of the residuals is  0.01

Tucker Lewis Index of factoring reliability =  0.995
RMSEA index =  0.023  and the 90 % confidence intervals are  0.022 0.025
BIC =  -66.89
```

Major and minor factors

- Real data are normally a mix of major factors plus many small factors.
- Variously called "correlated errors", "nuisance factors", "minor factors", real data are messy.
- Factor model: $\mathbf{R} = \mathbf{F}\phi\mathbf{F'} + \mathbf{mm'U}^2$
- These data can be simulated with `sim.minor`
  - Need to specify the major factor loadings, the minors default to have loadings = +/- .2
  - Also need to specify the number of variables
- Examine the various goodness of fit statistics as sample size increases

Simulating noisy data N = 100

```
> set.seed(42)
>  P <- sim.minor(9,3,fbig=c(.9,.8,.7),n=100)
> summary(fa(P$observed,3))
```

```
Factor analysis with Call: fa(r = P$observed, nfactors = 3)

Test of the hypothesis that 3 factors are sufficient.
The degrees of freedom for the model is 12  and the
       objective function was  0.07
The number of observations was  100  with
      Chi Square =  6.2  with prob <  0.91

The root mean square of the residuals is  0.02
The df corrected root mean square of the residuals is  0.04

Tucker Lewis Index of factoring reliability =  1.068
RMSEA index =  0  and the 90 % confidence intervals are  0 0.034
BIC =  -49.07
```

Simulating major and minor N = 400

```
> set.seed(42)
 P <- sim.minor(9,3,fbig=c(.9,.8,.7),n=400)
> P


Call: sim(fx = fload, n = n)

 $model (Population correlation matrix)
      V1    V2    V3    V4    V5    V6    V7    V8    V9
V1  1.00  0.49  0.63  0.04  0.00 -0.04  0.00 -0.04  0.00
V2  0.49  1.00  0.71 -0.04  0.00  0.00  0.04  0.00  0.00
V3  0.63  0.71  1.00 -0.04 -0.04  0.00  0.00 -0.04  0.00
V4  0.04 -0.04 -0.04  1.00 -0.49  0.55  0.00 -0.04  0.00
V5  0.00  0.00 -0.04 -0.49  1.00 -0.63  0.04  0.04  0.00
V6 -0.04  0.00  0.00  0.55 -0.63  1.00 -0.04  0.04  0.00
V7  0.00  0.04  0.00  0.00  0.04 -0.04  1.00 -0.45  0.63
V8 -0.04  0.00 -0.04 -0.04  0.04  0.04 -0.45  1.00 -0.63
V9  0.00  0.00  0.00  0.00  0.00  0.00  0.63 -0.63  1.00

$reliability (population reliability)
  V1   V2   V3   V4   V5   V6   V7   V8   V9
0.53 0.57 0.93 0.57 0.53 0.93 0.57 0.57 0.81
```

Exploratory structure N = 400

```
> f3 <- fa(P$observed,3)
> summary(f3)
```

```
Factor analysis with Call: fa(r = P$observed, nfactors = 3)

Test of the hypothesis that 3 factors are sufficient.
The degrees of freedom for the model is 12  and the
      objective function was  0.08
The number of observations was  400  with
       Chi Square =  32.91  with prob <  0.001

The root mean square of the residuals is  0.02
The df corrected root mean square of the residuals is  0.04

Tucker Lewis Index of factoring reliability =  0.949
RMSEA index =  0.067  and the 90 % confidence intervals are  0.066 0.07
BIC =  -38.98
```

N = 800 Major + minor factors

```
> set.seed(42)
>  P <- sim.minor(9,3,fbig=c(.9,.8,.7),n=800)
> summary(fa(P$observed,3))
```

```
Factor analysis with Call: fa(r = P$observed, nfactors = 3)

Test of the hypothesis that 3 factors are sufficient.
The degrees of freedom for the model is 12  and the
       objective function was  0.07
The number of observations was  800  with
       Chi Square =  54.05  with prob <  2.7e-07

The root mean square of the residuals is  0.02
The df corrected root mean square of the residuals is  0.04

Tucker Lewis Index of factoring reliability =  0.949
RMSEA index =  0.067  and the 90 % confidence intervals are  0.066 0.069
BIC =  -26.16
```

Confirmatory (Using lavaan)

```
> set.seed(42)
> P <- sim.minor(9,3,fbig=c(.9,.8,.7),n=100)
> my.data <- data.frame(P$observed)
> fit <- cfa(model,data =my.data,std.lv=TRUE)
> fit
```

```
>  set.seed(42)
>  P <- sim.minor(9,3,fbig=c(.9,.8,.7),n=100)
>  my.data <- data.frame(P$observed)
>  fit <- cfa(model,data =my.data,std.lv=TRUE)
> fit
Lavaan (0.4-5) converged normally after 28 iterations

  Number of observations                           100

  Estimator                                         ML
  Minimum Function Chi-square                   23.210
  Degrees of freedom                                24
  P-value                                        0.507
```

Lavaan N = 400

```
>  set.seed(42)
>  P <- sim.minor(9,3,fbig=c(.9,.8,.7),n=400)
>  my.data <- data.frame(P$observed)
>  fit <- cfa(model,data =my.data,std.lv=TRUE)
> fit

Lavaan (0.4-5) converged normally after 26 iterations

  Number of observations                           400

  Estimator                                         ML
  Minimum Function Chi-square                   45.293
  Degrees of freedom                                24
  P-value                                        0.005
```

Lavaan N = 800

```
>  set.seed(42)
>  P <- sim.minor(9,3,fbig=c(.9,.8,.7),n=800)
>  my.data <- data.frame(P$observed)
>  fit <- cfa(model,data =my.data,std.lv=TRUE)
> fit
```

```
Lavaan (0.4-5) converged normally after 28 iterations

  Number of observations                           800

  Estimator                                         ML
  Minimum Function Chi-square                   69.159
  Degrees of freedom                                24
  P-value                                        0.000
```

Lavaan N = 1600

```
> set.seed(42)
> P <- sim.minor(9,3,fbig=c(.9,.8,.7),n=1600)
> my.data <- data.frame(P$observed)
> fit <- cfa(model,data =my.data,std.lv=TRUE)
> fit

Lavaan (0.4-5) converged normally after 27 iterations

  Number of observations                          1600

  Estimator                                         ML
  Minimum Function Chi-square                  162.602
  Degrees of freedom                                24
  P-value                                        0.000
```

With parameter estimates

```
                     Estimate  Std.err  Z-value  P(>|z|)
Latent variables:
  F1 =~
    V1                  0.576    0.048   12.051    0.000
    V2                  0.634    0.048   13.168    0.000
    V3                  0.964    0.050   19.447    0.000
  F2 =~
    V4                  0.776    0.048   16.176    0.000
    V5                 -0.788    0.048  -16.489    0.000
    V6                  0.812    0.050   16.379    0.000
  F3 =~
    V7                  0.621    0.047   13.282    0.000
    V8                 -0.753    0.054  -14.017    0.000
    V9                  0.911    0.048   18.838    0.000

Covariances:
  F1 ~~
    F2                  0.066    0.055    1.185    0.236
    F3                  0.004    0.054    0.070    0.945
  F2 ~~
    F3                  0.011    0.059    0.188    0.851

Variances:
    V1                  0.568    0.047   12.072    0.000
    V2                  0.507    0.046   10.902    0.000
   ...
    V9                  0.142    0.057    2.494    0.013
    F1                  1.000
    F2                  1.000
    F3                  1.000
```

Why do models not fit?

- They are wrong
- The data are more complicated than we thought
  - This is the example of multiple minor factors. Not really part of our model, but there none the less
  - Perhaps the correlated residuals are because of something about the item, or about the way we collect the data.
- Because the data are in fact not normal.
  - Items are discrete, not continuous
  - This affects covariances.
- Once again, lets see this by simulation.

Simulate items

- First think of items as continuos variables
  - Simulate a two dimensional structure
  - Either a simple structure or a circumplex
- Then make the items discrete categories.
- Compare goodness of fits.
- Use the `sim.items` function

simulate 8 circumplex items, N = 500

```
> set.seed(42)
> X <- sim.item(8,circum=TRUE)
> str(X)   #what is the structure of the output?
> round(cor(X),2)
```

```
 num [1:500, 1:8] 0.136 0.44 -0.791 0.747 -1.068 ...
 - attr(*, "dimnames")=List of 2
  ..$ : NULL
  ..$ : chr [1:8] "V1" "V2" "V3" "V4" ...

> round(cor(X),2)
      V1    V2    V3    V4    V5    V6    V7    V8
V1  1.00  0.25  0.01 -0.27 -0.30 -0.21  0.05  0.24
V2  0.25  1.00  0.22 -0.06 -0.26 -0.29 -0.22  0.11
V3  0.01  0.22  1.00  0.20 -0.02 -0.28 -0.40 -0.23
V4 -0.27 -0.06  0.20  1.00  0.25  0.03 -0.32 -0.34
V5 -0.30 -0.26 -0.02  0.25  1.00  0.26 -0.01 -0.24
V6 -0.21 -0.29 -0.28  0.03  0.26  1.00  0.24  0.07
V7  0.05 -0.22 -0.40 -0.32 -0.01  0.24  1.00  0.21
V8  0.24  0.11 -0.23 -0.34 -0.24  0.07  0.21  1.00>
```

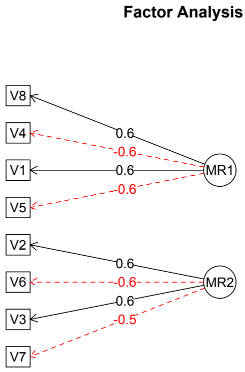
2 factors

```
> f2 <- fa(X,2)
> diagram(f2)
> diagram(f2,cut=.1)
> plot(f2)
> f2

Factor Analysis using method =  minres
Factor Analysis using method =  minres
Call: fa(r = X, nfactors = 2)
Standardized loadings based upon correlation matrix
     MR1   MR2   h2   u2
V1  0.01  0.55 0.31 0.69
V2 -0.35  0.45 0.32 0.68
V3 -0.61  0.04 0.37 0.63
V4 -0.40 -0.47 0.39 0.61
V5  0.04 -0.58 0.33 0.67
V6  0.43 -0.37 0.31 0.69
V7  0.65  0.05 0.43 0.57
V8  0.35  0.41 0.30 0.70
                MR1  MR2
SS loadings    1.38 1.38
Proportion Var 0.17 0.17
Cumulative Var 0.17 0.35
```

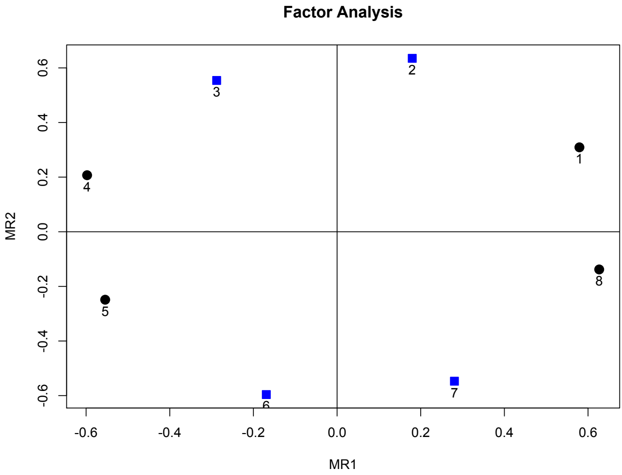
2 factors goodness of fit

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

The degrees of freedom for the null model are  28  and the objective function
          was  1.16 with Chi Square of  574.32
The degrees of freedom for the model are 13  and the
        objective function was  0.04

The root mean square of the residuals is  0.02
The df corrected root mean square of the residuals is  0.04
The number of observations was  500  with
       Chi Square =  19.21  with prob <  0.12

Tucker Lewis Index of factoring reliability =  0.975
RMSEA index =  0.031  and the 90 % confidence intervals are  0.031 0.036
BIC =  -61.58
Fit based upon off diagonal values = 0.99
Measures of factor score adequacy
                                                 MR1  MR2
Correlation of scores with factors              0.83 0.82
Multiple R square of scores with factors        0.69 0.67
```

Show the factor structure



Plot the items in the factor space



Simulate discrete items

```
> set.seed(42)
> X <- sim.item(8,500,TRUE,categorical=TRUE)
> round(cor(X),2)

      V1    V2    V3    V4    V5    V6    V7    V8
V1  1.00  0.26  0.02 -0.26 -0.27 -0.19  0.03  0.22
V2  0.26  1.00  0.21 -0.06 -0.23 -0.28 -0.20  0.12
V3  0.02  0.21  1.00  0.19 -0.03 -0.26 -0.36 -0.21
V4 -0.26 -0.06  0.19  1.00  0.21  0.01 -0.29 -0.35
V5 -0.27 -0.23 -0.03  0.21  1.00  0.28 -0.01 -0.21
V6 -0.19 -0.28 -0.26  0.01  0.28  1.00  0.19  0.08
V7  0.03 -0.20 -0.36 -0.29 -0.01  0.19  1.00  0.19
V8  0.22  0.12 -0.21 -0.35 -0.21  0.08  0.19  1.00
>
```

Exploratory factoring

```
> f2
Factor Analysis using method =  minres
Call: fa(r = X, nfactors = 2)
Standardized loadings based upon correlation matrix
     MR1   MR2   h2   u2
V1   0.24  0.48 0.29 0.71
V2  -0.10  0.54 0.31 0.69
V3  -0.51  0.30 0.35 0.65
V4  -0.58 -0.23 0.39 0.61
V5  -0.19 -0.50 0.29 0.71
V6   0.21 -0.51 0.30 0.70
V7   0.55 -0.21 0.35 0.65
V8   0.51  0.21 0.30 0.70
                MR1  MR2
SS loadings    1.31 1.27
Proportion Var 0.16 0.16
Cumulative Var 0.16 0.32
With factor correlations of
    MR1 MR2
MR1   1   0
MR2   0   1
```

Goodness of fit statistics

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

The degrees of freedom for the null model are  28  and the objective function w
The degrees of freedom for the model are 13  and the objective function was  0.

The root mean square of the residuals is  0.02
The df corrected root mean square of the residuals is  0.04
The number of observations was  500  with
       Chi Square =  24.74  with prob <  0.025

Tucker Lewis Index of factoring reliability =  0.947
RMSEA index =  0.043  and the 90 % confidence intervals are  0.043 0.046
BIC =  -56.05
Fit based upon off diagonal values = 0.98
Measures of factor score adequacy
                                                 MR1  MR2
Correlation of scores with factors              0.82 0.80
Multiple R square of scores with factors        0.67 0.65
Minimum correlation of possible factor scores  0.33 0.29
```

CFA of circumplex data – continuous case

```
>  set.seed(42)
> X <- sim.item(8,500,TRUE,categorical=FALSE)
> mod.circ <- 'F1 =~ V8 +V4  +V1  +V5
+                  F2 =~ V2 + V6 + V3 + V7'
> fit.circ.c <- cfa(mod.circ,data=X.df,std.lv=TRUE)
> summary(fit.circ.c)

avaan (0.4-5) converged normally after 31 iterations

  Number of observations                           500

  Estimator                                         ML
  Minimum Function Chi-square                  188.657
  Degrees of freedom                                19
  P-value                                        0.000

Parameter estimates:
```

CFA with lavaan of discrete items

```
> mod.circ <- 'F1 =~ V8 +V4  +V1  +V5
+                   F2 =~ V2 + V6 + V3 + V7'
> X.df <- data.frame(X)
> fit.circ <- cfa(mod.circ,data=X.df,std.lv=TRUE)
> fit.circ
```

```
Lavaan (0.4-5) converged normally after 31 iterations

  Number of observations                             500

  Estimator                                           ML
  Minimum Function Chi-square                    188.657
  Degrees of freedom                                  19
  P-value                                          0.000
```

cfa parameters

```
                  Estimate  Std.err  Z-value  P(>|z|)
Latent variables:
  F1 =~
    V8               0.589    0.062    9.548    0.000
    V4              -0.659    0.064  -10.283    0.000
    V1               0.429    0.059    7.238    0.000
    V5              -0.375    0.059   -6.317    0.000
  F2 =~
    V2               0.339    0.057    5.997    0.000
    V6              -0.398    0.059   -6.757    0.000
    V3               0.709    0.069   10.214    0.000
    V7              -0.609    0.064   -9.453    0.000

Covariances:
  F1 ~~
    F2              -0.268    0.071   -3.757    0.000

Variances:
    V8               0.712    0.070   10.224    0.000
    V4               0.622    0.076    8.188    0.000
    V1               0.910    0.066   13.692    0.000
```