HW 4

2019150432 임효진

May 27, 2021

1

```r
library(psych)
```

```
## Warning: package 'psych' was built under R version 3.6.3
```

```r
cor=matrix(rep(1, 36), nrow = 6)
cor[lower.tri(cor)]=c(0.505, 0.569, 0.602, 0.621, 0.603,
                      0.422, 0.467, 0.482, 0.450, 0.926,
                      0.877, 0.878, 0.874, 0.894, 0.937)
cor[upper.tri(cor, diag=F)]=t(cor)[upper.tri(cor, diag = F)]

facMl=fac(cor, nfactors = 2, fm="ml", rotate = "none")
facMl
```

```
## Factor Analysis using method =  ml
## Call: fac(r = cor, nfactors = 2, rotate = "none", fm = "ml")
## Standardized loadings (pattern matrix) based upon correlation matrix
##    ML1   ML2   h2     u2 com
## 1 0.63  0.06 0.40 0.5981 1.0
## 2 0.49  0.05 0.24 0.7586 1.0
## 3 0.94 -0.07 0.88 0.1204 1.0
## 4 0.97 -0.22 1.00 0.0050 1.1
## 5 0.96  0.27 0.99 0.0062 1.2
## 6 0.94  0.11 0.91 0.0949 1.0
##
##                        ML1  ML2
## SS loadings           4.27 0.15
## Proportion Var        0.71 0.02
## Cumulative Var        0.71 0.74
## Proportion Explained  0.97 0.03
## Cumulative Proportion 0.97 1.00
##
## Mean item complexity =  1.1
## Test of the hypothesis that 2 factors are sufficient.
##
## The degrees of freedom for the null model are  15  and the objective function was  6.75
```

```
## The degrees of freedom for the model are 4  and the objective function was  0.1
##
## The root mean square of the residuals (RMSR) is  0.05
## The df corrected root mean square of the residuals is  0.1
##
## Fit based upon off diagonal values = 0.99
## Measures of factor score adequacy
##                                                    ML1  ML2
## Correlation of (regression) scores with factors   1.00 0.98
## Multiple R square of scores with factors          1.00 0.96
## Minimum correlation of possible factor scores     0.99 0.91
```

(a)

```
facMl$communalities
```

```
##        V1        V2        V3        V4        V5        V6
## 0.4019249 0.2414293 0.8796473 0.9950000 0.9937647 0.9051294
```

(b)

```
facMl$uniquenesses
```

```
## [1] 0.598067467 0.758574445 0.120353359 0.004999065 0.006235315 0.094871844
```

(c)

```
facMl$loadings
```

```
##
## Loadings:
##      ML1    ML2
## [1,]  0.631
## [2,]  0.489
## [3,]  0.935
## [4,]  0.973 -0.218
## [5,]  0.959  0.272
## [6,]  0.944  0.115
##
##                  ML1   ML2
## SS loadings    4.271 0.146
## Proportion Var 0.712 0.024
## Cumulative Var 0.712 0.736
```

The empty space implies that the loadings are less than 0.1. From the loadings above, we can interpret the correlation between skull breadth and F2 is less than 0.1.

(d)

```r
Lambda <- facMl$loadings
Psi <- diag(facMl$uniquenesses)
S <- facMl$r
Sigma <- Lambda %*% t(Lambda) + Psi
round(S - Sigma, 6)
```

```
##           [,1]      [,2]      [,3]      [,4]      [,5]      [,6]
## [1,]  0.000000  0.193493 -0.017096  0.000243 -0.000262  0.000101
## [2,]  0.193493  0.000000 -0.032149  0.000799  0.000403 -0.017277
## [3,] -0.017096 -0.032149  0.000000  0.000022  0.000013  0.003270
## [4,]  0.000243  0.000799  0.000022  0.000000  0.000000 -0.000072
## [5,] -0.000262  0.000403  0.000013  0.000000  0.000000 -0.000003
## [6,]  0.000101 -0.017277  0.003270 -0.000072 -0.000003  0.000000
```

The numbers are close to 0 which means that the factor model is a good representation of the manifest variable.

- Repeat with Varimax Rotation

```r
facM2=fac(cor, nfactors = 2, fm="ml", rotate = "varimax")
facM2
```

```
## Factor Analysis using method =  ml
## Call: fac(r = cor, nfactors = 2, rotate = "varimax", fm = "ml")
## Standardized loadings (pattern matrix) based upon correlation matrix
##    ML2  ML1   h2     u2 com
## 1 0.48 0.41 0.40 0.5981 2.0
## 2 0.37 0.32 0.24 0.7586 1.9
## 3 0.60 0.72 0.88 0.1204 1.9
## 4 0.52 0.85 1.00 0.0050 1.7
## 5 0.86 0.50 0.99 0.0062 1.6
## 6 0.74 0.60 0.91 0.0949 1.9
##
##                        ML2  ML1
## SS loadings           2.31 2.11
## Proportion Var        0.38 0.35
## Cumulative Var        0.38 0.74
## Proportion Explained  0.52 0.48
## Cumulative Proportion 0.52 1.00
##
## Mean item complexity =  1.8
## Test of the hypothesis that 2 factors are sufficient.
##
## The degrees of freedom for the null model are  15  and the objective function was  6.75
## The degrees of freedom for the model are 4  and the objective function was  0.1
##
## The root mean square of the residuals (RMSR) is  0.05
## The df corrected root mean square of the residuals is  0.1
```

```
##
## Fit based upon off diagonal values = 0.99
## Measures of factor score adequacy
##                                                   ML2  ML1
## Correlation of (regression) scores with factors   0.99 0.99
## Multiple R square of scores with factors          0.98 0.98
## Minimum correlation of possible factor scores     0.95 0.95
```

```
facM2$communalities
```

```
##        V1        V2        V3        V4        V5        V6
## 0.4019249 0.2414293 0.8796473 0.9950000 0.9937647 0.9051294
```

```
facM2$uniquenesses
```

```
## [1] 0.598067467 0.758574445 0.120353359 0.004999065 0.006235315 0.094871844
```

```
facM2$loadings
```

```
##
## Loadings:
##      ML2   ML1
## [1,] 0.483 0.411
## [2,] 0.375 0.318
## [3,] 0.601 0.720
## [4,] 0.524 0.849
## [5,] 0.865 0.496
## [6,] 0.742 0.595
##
##                  ML2   ML1
## SS loadings    2.308 2.109
## Proportion Var 0.385 0.351
## Cumulative Var 0.385 0.736
```

From the loadings above, we can interpret the correlation between skull breadth and F2 is 0.375.

```
Lambda <- facM2$loadings
Psi <- diag(facM2$uniquenesses)
S <- facM2$r
Sigma <- Lambda %*% t(Lambda) + Psi
round(S - Sigma, 6)
```

```
##           [,1]      [,2]      [,3]      [,4]      [,5]      [,6]
## [1,]  0.000000  0.193493 -0.017096  0.000243 -0.000262  0.000101
## [2,]  0.193493  0.000000 -0.032149  0.000799  0.000403 -0.017277
## [3,] -0.017096 -0.032149  0.000000  0.000022  0.000013  0.003270
## [4,]  0.000243  0.000799  0.000022  0.000000  0.000000 -0.000072
## [5,] -0.000262  0.000403  0.000013  0.000000  0.000000 -0.000003
## [6,]  0.000101 -0.017277  0.003270 -0.000072 -0.000003  0.000000
```

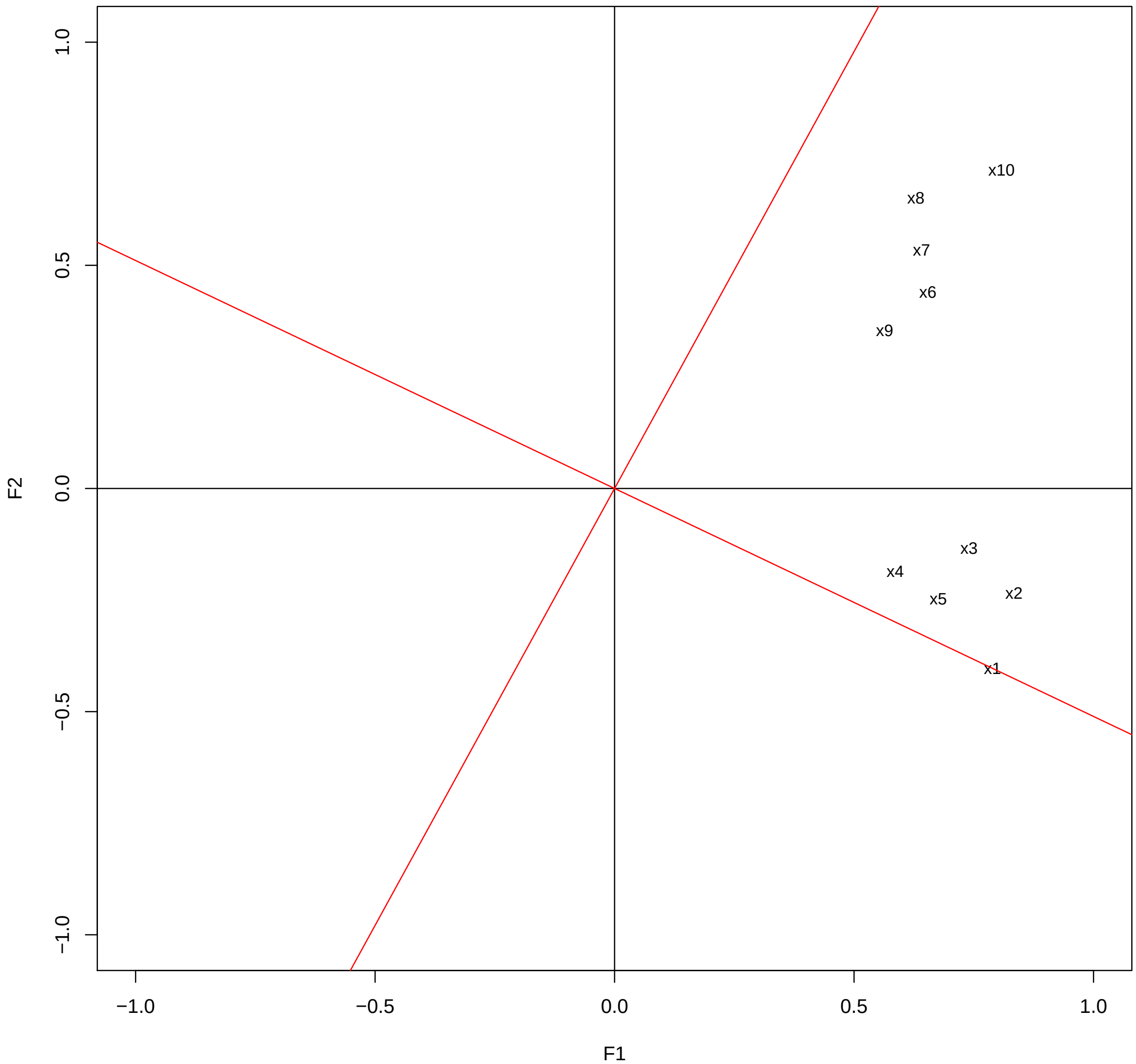
The residual matrix calculated with varimax rotation is identical with the former one.

2

(a)

```r
f1=c(0.789, 0.834, 0.740, 0.586, 0.676, 0.654, 0.641,
     0.629, 0.564, 0.808)
f2=c(-0.403, -0.234, -0.134, -0.185, -0.248, 0.440, 0.534,
     0.651, 0.354, 0.714)
f=cbind(f1, f2)
rownames(f)=c("x1", "x2", "x3", "x4", "x5",
              "x6", "x7", "x8", "x9", "x10")

plot(f1, f2, xlab="F1", ylab="F2", xlim=c(-1, 1),
     ylim=c(-1, 1), type="n")
text(f1, f2, labels = rownames(f), cex=0.8)
abline(h=0, v=0)
abline(coef=c(0, -0.403/0.789), col="red")
abline(coef=c(0, 0.789/0.403), col="red")
```




```r
# angle of rotation
atan2(0.403, 0.789)*180/pi
```

```
## [1] 27.05672
```

When using the coordinates of x1 to calculate the angle of roatation, the minimum angle needed to remove those that are negatve is approximately 27.057.

(b)

From the above graph we can group x1, x2, x3, x4, x5 associated with factor 1 and the others associated with factor 2.

3

```
cor=matrix(c(1, 1/3, 1/3, 1/3, 1, 1/10, 1/3, 1/10, 1),
           nrow = 3, byrow = T)

fac1=fac(cor, nfactors = 1, fm="ml", rotate = "none")
fac1
```

```
## Factor Analysis using method =  ml
## Call: fac(r = cor, nfactors = 1, rotate = "none", fm = "ml")
## Standardized loadings (pattern matrix) based upon correlation matrix
##    ML1   h2    u2 com
## 1 1.00 1.00 0.005   1
## 2 0.33 0.11 0.888   1
## 3 0.33 0.11 0.888   1
##
##                 ML1
## SS loadings    1.22
## Proportion Var 0.41
##
## Mean item complexity =  1
## Test of the hypothesis that 1 factor is sufficient.
##
## The degrees of freedom for the null model are  3  and the objective function was  0.24
## The degrees of freedom for the model are 0  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  NA
##
## Fit based upon off diagonal values = 1
## Measures of factor score adequacy
##                                                  ML1
## Correlation of (regression) scores with factors  1.00
## Multiple R square of scores with factors         1.00
## Minimum correlation of possible factor scores    0.99
```

The specific variances of the second and third vairable are significantly large. The variance implies that about 89% of the variability can not be explained by the factors. Therefore using only one factor is not adequate.

4

(a)

```
sales=read.table("sales.dat", header = T)
cor=cor(sales)
```

```r
fac2=fac(cor, nfactors = 2, fm="ml", rotate = "varimax")

# loadings
fac2$loadings
```

```
##
## Loadings:
##    ML2   ML1
## X1 0.852 0.452
## X2 0.868 0.419
## X3 0.717 0.602
## X4 0.148 0.987
## X5 0.501 0.525
## X6 0.619
## X7 0.946 0.277
##
##                  ML2   ML1
## SS loadings    3.545 2.071
## Proportion Var 0.506 0.296
## Cumulative Var 0.506 0.802
```

```r
# communalities
fac2$communalities
```

```
##        X1        X2        X3        X4        X5        X6        X7
## 0.9308084 0.9296171 0.8766888 0.9950000 0.5264156 0.3863585 0.9711829
```

```r
# specific variance
fac2$uniquenesses
```

```
##          X1          X2          X3          X4          X5          X6
## 0.069191515 0.070382519 0.123310305 0.004987889 0.473588339 0.613637151
##          X7
## 0.028817227
```

```r
# sigma
Lambda2=fac2$loadings
Psi2=diag(fac2$uniquenesses)
Sigma2=Lambda2 %*% t(Lambda2) + Psi2

Sigma2
```

```
##           X1        X2        X3        X4        X5        X6        X7
## X1 1.0000000 0.9295185 0.8834713 0.5720627 0.6642314 0.5543376 0.9311883
## X2 0.9295185 1.0000000 0.8749727 0.5413952 0.6547844 0.5624010 0.9373106
## X3 0.8834713 0.8749727 1.0000000 0.6996404 0.6751659 0.4798313 0.8449575
## X4 0.5720627 0.5413952 0.6996404 1.0000000 0.5918666 0.1504776 0.4126431
## X5 0.6642314 0.6547844 0.6751659 0.5918666 1.0000000 0.3412918 0.6189365
## X6 0.5543376 0.5624010 0.4798313 0.1504776 0.3412918 1.0000000 0.6017606
## X7 0.9311883 0.9373106 0.8449575 0.4126431 0.6189365 0.6017606 1.0000000
```

From the loadings we can suggest that variables can be grouped by x4, x5 with factor 1 and the others with factor 2. Thus factor 1 implies ability of creativity and factor 2 implies sales skills.

(b)

```
fac3=fac(cor, nfactors = 3, fm="ml", rotate = "varimax")

# loadings
fac3$loadings
```

```
##
## Loadings:
##    ML3   ML1   ML2
## X1 0.793 0.374 0.438
## X2 0.911 0.317 0.185
## X3 0.651 0.544 0.438
## X4 0.255 0.964
## X5 0.542 0.465 0.207
## X6 0.299       0.950
## X7 0.917 0.180 0.298
##
##                  ML3   ML1   ML2
## SS loadings    3.175 1.718 1.453
## Proportion Var 0.454 0.245 0.208
## Cumulative Var 0.454 0.699 0.906
```

```
# communalities
fac3$communalities
```

```
##        X1        X2        X3        X4        X5        X6        X7
## 0.9614288 0.9655182 0.9118758 0.9950000 0.5533880 0.9950000 0.9624919
```

```
# specific variance
fac3$uniquenesses
```

```
##          X1          X2          X3          X4          X5          X6
## 0.038571357 0.034481797 0.088124265 0.004956641 0.446619476 0.004968310
##          X7
## 0.037508685
```

```
# sigma
Lambda3=fac3$loadings
Psi3=diag(fac3$uniquenesses)
Sigma3=Lambda3 %*% t(Lambda3) + Psi3

Sigma3
```

```
##           X1        X2        X3        X4        X5        X6        X7
## X1 1.0000000 0.9228132 0.9120900 0.5714373 0.6949346 0.6738835 0.9255325
## X2 0.9228132 1.0000000 0.8471023 0.5417814 0.6799445 0.4654561 0.9481931
```

```
## X3 0.9120900 0.8471023 1.0000000 0.6991263 0.6969684 0.6402870 0.8255831
## X4 0.5714373 0.5417814 0.6991263 1.0000000 0.5910467 0.1469508 0.4130097
## X5 0.6949346 0.6799445 0.6969684 0.5910467 1.0000000 0.3841949 0.6425632
## X6 0.6738835 0.4654561 0.6402870 0.1469508 0.3841949 1.0000000 0.5669006
## X7 0.9255325 0.9481931 0.8255831 0.4130097 0.6425632 0.5669006 1.0000000
```

From the loadins we can suggest that variables can be grouped by x4 with factor 1, x6 with factor 2 and the others with factor 3. Thus factor 1 implies creativity, factor 2 implies abstract reasoning and factor 3 implies sales skills.

(c)

```r
# residual matrix of the first model
round(cor - Sigma2, 6)
```

```
##           X1        X2        X3        X4        X5        X6        X7
## X1  0.000000 -0.003443  0.000531 -0.000026  0.043842  0.120070 -0.003877
## X2 -0.003443  0.000000 -0.032449  0.000113  0.091125 -0.097013  0.006985
## X3  0.000531 -0.032449  0.000000  0.000723 -0.037695  0.161257  0.007611
## X4 -0.000026  0.000113  0.000723  0.000000 -0.001131 -0.003570 -0.000004
## X5  0.043842  0.091125 -0.037695 -0.001131  0.000000  0.044658 -0.044383
## X6  0.120070 -0.097013  0.161257 -0.003570  0.044658  0.000000 -0.035388
## X7 -0.003877  0.006985  0.007611 -0.000004 -0.044383 -0.035388  0.000000
```

```r
# residual matrix of the second model
round(cor - Sigma3, 6)
```

```
##           X1        X2        X3        X4        X5        X6        X7
## X1  0.000000  0.003263 -0.028088  0.000599  0.013139  0.000524  0.001779
## X2  0.003263  0.000000 -0.004579 -0.000273  0.065965 -0.000068 -0.003897
## X3 -0.028088 -0.004579  0.000000  0.001237 -0.059497  0.000802  0.026985
## X4  0.000599 -0.000273  0.001237  0.000000 -0.000311 -0.000043 -0.000370
## X5  0.013139  0.065965 -0.059497 -0.000311  0.000000  0.001755 -0.068010
## X6  0.000524 -0.000068  0.000802 -0.000043  0.001755  0.000000 -0.000528
## X7  0.001779 -0.003897  0.026985 -0.000370 -0.068010 -0.000528  0.000000
```

The model using three-factor solution has a residual matrix close to 0 compared to the one with two-factor solution. Therefore three factor seems to be an adequate choice.

(d)

```r
fac2$dof
```

```
## [1] 8
```

```r
fac2$objective
```

```
## [1] 2.633716
```

```r
pchisq(2.63, 8)
```

```
## [1] 0.04460701
```

```r
fac3$dof
```

```
## [1] 3
```

```r
fac3$objective
```

```
## [1] 1.418449
```

```r
pchisq(1.42, 3)
```

```
## [1] 0.2991466
```

$U = n'min(F)$ is not significant when 3-factor model is chosen which implies that using 3 factors is more adequate.

(e)

```r
factor.scores(sales, fac2)$scores
```

```
##               ML2         ML1
##  [1,] -0.86288507 -0.44207987
##  [2,] -1.17538519 -0.90503168
##  [3,] -0.22205438 -0.79478486
##  [4,] -0.13640449  0.48423974
##  [5,]  0.40665543 -0.36962803
##  [6,] -0.73944040 -0.20034064
##  [7,] -0.39319904 -0.51383740
##  [8,]  1.51969565  1.50942456
##  [9,]  0.41697990 -0.36442007
## [10,]  0.78697654  0.58616049
## [11,]  0.29924822  0.16051109
## [12,]  0.28370806 -0.35154454
## [13,]  0.09445898  1.21144685
## [14,]  0.57239264 -0.91541423
## [15,]  0.05433039  0.43356990
## [16,] -1.38344667 -0.88556211
## [17,]  0.19428237 -0.08967132
## [18,]  0.57427848 -0.14303420
## [19,]  0.56521867 -1.66709766
## [20,] -0.71830309  1.58950715
## [21,] -1.47489872 -0.09735094
## [22,]  1.61917360 -1.83569672
## [23,] -1.41138793 -0.36781786
## [24,]  0.76202605  0.08769452
## [25,]  0.73786158  1.12102954
## [26,] -0.65995849 -0.21671643
```

```
## [27,]  0.83749784  0.58969159
## [28,]  1.85029030 -0.59941450
## [29,] -0.99700848 -0.67250456
## [30,]  1.67045843 -0.81609158
## [31,]  0.68466173  1.62758716
## [32,] -2.13446067  0.77021129
## [33,]  0.10438389 -1.09015789
## [34,] -0.98294920 -0.16633005
## [35,]  0.69360463  1.63189252
## [36,]  1.90173183 -1.11169605
## [37,] -1.79699467  2.00312533
## [38,] -0.32209868  0.49887134
## [39,]  1.02249669  0.82203947
## [40,]  0.75013750  0.60996836
## [41,]  0.59843962 -0.64658417
## [42,] -0.98087780  0.59557190
## [43,] -0.12474833  1.50464803
## [44,] -0.82300552 -2.49193247
## [45,] -0.15031821 -1.04937759
## [46,]  0.10769164  1.71588510
## [47,] -1.07901129 -0.90856942
## [48,] -1.89412470 -0.54669082
## [49,]  0.36568659  0.65226431
## [50,]  0.98859381  0.05403743
```