Quant_II_hwk_03

吴温泉

目录

1   因子分析 (Factor Analysis)

对 abs.dta 数据中各个机构、个人信任程度的差异进行因子分析进行降维处理。

1.1   建立各个信任变量的描述性统计表格。

```
library(haven)
options(digits=2)


setwd("E:/SynologyDrive/Github/Quantitative_Analysis_II/hwk03/")
data <- read_dta('./abs.dta')
summary(data)
```

```
##     trustCourt     trustNatGov     trustParty    trustParliament trustCivService
## Min.    :1      Min.    :1      Min.    :1      Min.    :1       Min.    :1
## 1st Qu.:3      1st Qu.:3      1st Qu.:3      1st Qu.:3       1st Qu.:2
## Median :3      Median :4      Median :4      Median :4       Median :2
## Mean    :3      Mean    :3      Mean    :4      Mean    :3       Mean    :3
## 3rd Qu.:3      3rd Qu.:4      3rd Qu.:4      3rd Qu.:4       3rd Qu.:3
## Max.    :4      Max.    :4      Max.    :4      Max.    :4       Max.    :4
## NA's    :325   NA's    :136   NA's    :127   NA's    :242    NA's    :163
##   trustMilitary  trustPolice    trustLocGov    trustNewspaper     trustTV
## Min.    :1      Min.    :1      Min.    :1      Min.    :1       Min.    :1
## 1st Qu.:3      1st Qu.:3      1st Qu.:3      1st Qu.:3       1st Qu.:3
## Median :4      Median :3      Median :3      Median :3       Median :3
## Mean    :3      Mean    :3      Mean    :3      Mean    :3       Mean    :3
## 3rd Qu.:4      3rd Qu.:4      3rd Qu.:3      3rd Qu.:3       3rd Qu.:3
## Max.    :4      Max.    :4      Max.    :4      Max.    :4       Max.    :4
## NA's    :152   NA's    :155   NA's    :193   NA's    :419    NA's    :203
##     trustNGO     trustRelative  trustNeighbor   trustOther
## Min.    :1      Min.    :1      Min.    :1      Min.    :1
## 1st Qu.:2      1st Qu.:3      1st Qu.:3      1st Qu.:2
## Median :2      Median :3      Median :3      Median :3
## Mean    :2      Mean    :3      Mean    :3      Mean    :3
## 3rd Qu.:3      3rd Qu.:4      3rd Qu.:3      3rd Qu.:3
## Max.    :4      Max.    :4      Max.    :4      Max.    :4
## NA's    :750   NA's    :52    NA's    :72     NA's    :204
```

```
data <- na.omit(data)
```

1.2 观察各个信任变量间的相关性矩阵，你是否可以总结出初步的分类？请使用 corrplot()。

```
n <- dim(data)[1]
corMat <- cor(data, use = "pairwise.complete.obs") #correlation matrix
library(corrplot)
```

```
## Warning: package 'corrplot' was built under R version 4.0.3
```

```
## corrplot 0.84 loaded
```

```
res <- cor.mtest(data, conf.level = 0.95)
corrplot(corMat, method = "ellipse", type = "lower", p.mat = res$p, sig.level = 0.05, order = "hcl
```


可以发现：

1. trustNGO 是唯一与其他各项呈现**负相关**的一项，其中与 trustNewspaper、trustTV 的负相关系数最大。
2. trustRelative、trustNeighbor、trustOther、trustCivService 四项**个人信任**的变量相互之间的相关性较小，与其他信任变量的相关性也较小。
3. trustNewspaper 与 trustTV 两项**媒体信任**的相关性很大

4. trustMilitary、trustParliament、rustNatGov、trustParty、trustPolice、trustCourt、trustLocGov 七
   项**机构信任**变量之间的相关程度较大。其中可细分为：

   - trustMilitary、trustParliament、rustNatGov、trustParty 四项**党国信任**变量之间的相关程度很
     高
   - trustCourt **法院信任**与其他各项政府信任之间的相关性较低
   - rustNatGov **当地政府信任**与与其他各项政府信任之间的相关性较低，但与 trustPolice 之间的信
     任程度较高。

1.3   进行克隆巴赫系数（Cronbach's ）检验信任程度问题的内在信度（提示：是否都是同向？如果不是，进行必要的调整）。

trustNGO 与其他变量不同向，进行调整

```
library(psych)
```

```
## Warning: package 'psych' was built under R version 4.0.3
```

```
alpha(data, check.keys=TRUE)
```

```
## Warning in alpha(data, check.keys = TRUE): Some items were negatively correlated with total sca
##  This is indicated by a negative sign for the variable name.
```

```
##
## Reliability analysis
## Call: alpha(x = data, check.keys = TRUE)
##
##   raw_alpha std.alpha G6(smc) average_r S/N   ase mean   sd median_r
##       0.87      0.87     0.9      0.32 6.6 0.004  3.1 0.38     0.29
##
##  lower alpha upper     95% confidence boundaries
## 0.86 0.87 0.87
##
##  Reliability if an item is dropped:
##               raw_alpha std.alpha G6(smc) average_r S/N alpha se var.r med.r
## trustCourt         0.85      0.86    0.89      0.32 6.0   0.0044 0.030  0.28
## trustNatGov        0.85      0.85    0.88      0.31 5.9   0.0045 0.023  0.28
## trustParty         0.85      0.85    0.88      0.31 5.9   0.0045 0.023  0.28
## trustParliament    0.85      0.85    0.88      0.31 5.8   0.0045 0.023  0.28
```

```
## trustCivService        0.87          0.87          0.90          0.35 6.9    0.0038 0.028  0.31
## trustMilitary          0.85          0.86          0.89          0.31 6.0    0.0044 0.026  0.28
## trustPolice            0.85          0.85          0.88          0.31 5.8    0.0046 0.028  0.26
## trustLocGov            0.85          0.85          0.88          0.31 5.9    0.0045 0.029  0.26
## trustNewspaper         0.85          0.86          0.88          0.31 6.0    0.0044 0.029  0.26
## trustTV                0.85          0.86          0.88          0.32 6.0    0.0044 0.028  0.26
## trustNGO-              0.87          0.87          0.89          0.33 6.5    0.0041 0.029  0.30
## trustRelative          0.87          0.87          0.90          0.34 6.6    0.0041 0.030  0.31
## trustNeighbor          0.86          0.86          0.89          0.33 6.3    0.0042 0.031  0.30
## trustOther             0.87          0.87          0.90          0.34 6.7    0.0040 0.027  0.31
##
##   Item statistics
##                     n raw.r std.r r.cor r.drop mean   sd
## trustCourt       2310  0.66  0.65  0.62   0.58  3.1 0.64
## trustNatGov      2310  0.70  0.71  0.71   0.64  3.5 0.58
## trustParty       2310  0.69  0.70  0.70   0.62  3.5 0.63
## trustParliament 2310  0.71  0.71  0.71   0.65  3.4 0.61
## trustCivService 2310  0.41  0.39  0.30   0.28  2.5 0.73
## trustMilitary   2310  0.67  0.67  0.65   0.60  3.4 0.60
## trustPolice     2310  0.74  0.73  0.71   0.67  3.1 0.69
## trustLocGov     2310  0.71  0.70  0.68   0.64  3.0 0.68
## trustNewspaper  2310  0.67  0.67  0.66   0.60  2.9 0.63
## trustTV         2310  0.67  0.66  0.65   0.59  2.9 0.62
## trustNGO-       2310  0.53  0.51  0.46   0.42  2.6 0.72
## trustRelative   2310  0.43  0.45  0.38   0.34  3.3 0.52
## trustNeighbor   2310  0.52  0.55  0.50   0.45  3.0 0.44
## trustOther      2310  0.41  0.42  0.35   0.31  2.8 0.59
##
## Non missing response frequency for each item
##                    1    2    3    4 miss
## trustCourt      0.02 0.12 0.63 0.24    0
## trustNatGov     0.01 0.03 0.45 0.52    0
## trustParty      0.01 0.04 0.40 0.55    0
## trustParliament 0.01 0.04 0.45 0.50    0
## trustCivService 0.02 0.58 0.27 0.12    0
## trustMilitary   0.01 0.04 0.46 0.50    0
## trustPolice     0.02 0.14 0.56 0.28    0
## trustLocGov     0.02 0.17 0.61 0.20    0
## trustNewspaper  0.02 0.23 0.63 0.12    0
```

```
## trustTV          0.01 0.19 0.65 0.15    0
## trustNGO         0.08 0.52 0.33 0.06    0
## trustRelative    0.00 0.03 0.68 0.29    0
## trustNeighbor    0.00 0.09 0.82 0.09    0
## trustOther       0.02 0.23 0.69 0.06    0
```

可知总体的 Cronbach's 值为 0.8661063，通过检验。去掉某一个项变量的 Cronbach's 值中，raw_alpha 并没有明显变化，因此不进行去除。

1.4 进行主成分 (pcf) 因子分析，有多少因子的特征值 (Eigenvalue) 超过 1？这些因子的贡献度为何？需要多少因子，公共因子的累计贡献度才会超过 80%？是否存在特殊度 (Uniqueness) 过高的变量？观察从这几个统计量，你可以总结出几个因子来对机构信任程度进行降维分类？

```
# ev <- eigen(corMat)
# ev$values
#
# pcafit <- princomp(~., data=data, cor=TRUE)
# summary(pcafit) # print variance accounted for
# loadings(pcafit) # pc loadings
# plot(pcafit,type="lines") # scree plot
# # pcafit$scores # the principal components
# # biplot(pcafit)



# fit1 <- principal(dat, nfactors=14, rotate="varimax")

fit1 <- principal(data, nfactors = 14, rotate="none")
print(fit1, digits=2)
```

```
## Principal Components Analysis
## Call: principal(r = data, nfactors = 14, rotate = "none")
## Standardized loadings (pattern matrix) based upon correlation matrix
##                   PC1   PC2   PC3   PC4   PC5   PC6   PC7   PC8   PC9  PC10
## trustCourt        0.67  0.02 -0.08  0.05  0.45  0.12 -0.02  0.48 -0.28  0.04
## trustNatGov       0.75  0.45  0.09 -0.05 -0.19 -0.07 -0.01  0.09 -0.06 -0.15
## trustParty        0.75  0.46  0.09 -0.08 -0.20 -0.08 -0.05  0.02 -0.04 -0.18
## trustParliament   0.76  0.44  0.07 -0.06 -0.16 -0.08 -0.07  0.00  0.02 -0.11
```

```
## trustCivService  0.34 -0.08 -0.04  0.91 -0.09 -0.17  0.01  0.05  0.10 -0.01
## trustMilitary    0.71  0.39  0.04 -0.03 -0.07 -0.03 -0.07 -0.13  0.07  0.53
## trustPolice      0.75  0.00 -0.07 -0.02  0.45  0.01  0.00 -0.11  0.15  0.10
## trustLocGov      0.72 -0.04 -0.09  0.00  0.40  0.05  0.12 -0.37  0.11 -0.23
## trustNewspaper   0.65 -0.42 -0.38 -0.10 -0.26  0.10  0.20  0.08  0.05  0.04
## trustTV          0.65 -0.36 -0.44 -0.09 -0.25  0.11  0.21  0.07  0.08  0.00
## trustNGO        -0.46  0.55  0.24  0.03  0.08  0.10  0.60  0.14  0.18  0.03
## trustRelative    0.39 -0.26  0.56  0.08 -0.11  0.62 -0.15  0.04  0.18 -0.03
## trustNeighbor    0.49 -0.43  0.50  0.05 -0.08 -0.10  0.28 -0.21 -0.42  0.06
## trustOther       0.35 -0.51  0.44 -0.23  0.06 -0.47 -0.01  0.23  0.30 -0.01
##                  PC11  PC12  PC13  PC14 h2      u2 com
## trustCourt       0.13  0.02  0.01  0.02  1  4.4e-16 3.3
## trustNatGov     -0.07 -0.27  0.00 -0.26  1  2.1e-15 2.7
## trustParty      -0.07 -0.05  0.14  0.32  1  5.6e-16 2.7
## trustParliament  0.03  0.36 -0.18 -0.08  1  4.4e-16 2.5
## trustCivService  0.00  0.00  0.00  0.01  1 -2.2e-16 1.4
## trustMilitary    0.18 -0.05  0.03  0.01  1 -2.2e-16 2.8
## trustPolice     -0.44  0.02  0.00 -0.01  1 -4.4e-16 2.6
## trustLocGov      0.29 -0.06 -0.01  0.00  1 -6.7e-16 3.1
## trustNewspaper  -0.02 -0.12 -0.29  0.11  1  4.4e-16 4.0
## trustTV          0.01  0.13  0.29 -0.10  1  4.4e-16 4.0
## trustNGO         0.00  0.02 -0.01  0.02  1  0.0e+00 3.7
## trustRelative    0.01  0.00  0.01  0.00  1 -4.4e-16 3.6
## trustNeighbor   -0.06  0.04  0.01 -0.01  1 -2.2e-16 5.2
## trustOther       0.08 -0.01  0.01  0.00  1 -4.4e-16 5.5
##
##                        PC1  PC2  PC3  PC4  PC5  PC6  PC7  PC8  PC9 PC10 PC11
## SS loadings           5.42 1.88 1.19 0.92 0.84 0.72 0.57 0.54 0.47 0.42 0.35
## Proportion Var        0.39 0.13 0.09 0.07 0.06 0.05 0.04 0.04 0.03 0.03 0.02
## Cumulative Var        0.39 0.52 0.61 0.67 0.73 0.78 0.83 0.86 0.90 0.93 0.95
## Proportion Explained  0.39 0.13 0.09 0.07 0.06 0.05 0.04 0.04 0.03 0.03 0.02
## Cumulative Proportion 0.39 0.52 0.61 0.67 0.73 0.78 0.83 0.86 0.90 0.93 0.95
##                       PC12 PC13 PC14
## SS loadings           0.25 0.22 0.20
## Proportion Var        0.02 0.02 0.01
## Cumulative Var        0.97 0.99 1.00
## Proportion Explained  0.02 0.02 0.01
## Cumulative Proportion 0.97 0.99 1.00
##
```

```
## Mean item complexity =  3.4
## Test of the hypothesis that 14 components are sufficient.
##
## The root mean square of the residuals (RMSR) is  0
##  with the empirical chi square  0  with prob <  NA
##
## Fit based upon off diagonal values = 1
```

```
fit1$uniquenesses
```

```
##      trustCourt       trustNatGov        trustParty trustParliament trustCivService
##        4.4e-16          2.1e-15           5.6e-16         4.4e-16        -2.2e-16
##   trustMilitary        trustPolice       trustLocGov   trustNewspaper         trustTV
##       -2.2e-16         -4.4e-16          -6.7e-16         4.4e-16         4.4e-16
##        trustNGO      trustRelative    trustNeighbor       trustOther
##        0.0e+00         -4.4e-16          -2.2e-16        -4.4e-16
```

可知，PC1，PC2，PC3 三个因子的特征值超过了 1，贡献度为 0.3874151，0.1344018，0.08529483。需要七个因子，公共因子的累计贡献度才会超过 80%。没有特殊值过高的变量。PC3 对 trustCourt、trustNatGov、trustParty、trustParliament、trustMilitary、trustPolice、trustLocGov 的因子载荷较高，均大于 0。可以使用 PC1 对机构信任程度进行降维。

1.5 进行因子旋转。

```
M1 <- factanal(~., data = data, factors = 3 , scores = "regression")
M2 <- update(M1, factor = 3, roation = "varimax", SMC = T)
print(M2, digits=2, cutoff=.3, sort=TRUE)
```

```
##
## Call:
## factanal(x = ~., factors = 3, data = data, scores = "regression",     factor = 3, roation = "va
##
## Uniquenesses:
##      trustCourt       trustNatGov        trustParty trustParliament trustCivService
##          0.64             0.24              0.22            0.24             0.91
##   trustMilitary        trustPolice       trustLocGov   trustNewspaper         trustTV
##          0.45             0.53              0.56            0.25             0.21
##        trustNGO      trustRelative    trustNeighbor       trustOther
```

```
##              0.64            0.76            0.54            0.66
##
## Loadings:
##                Factor1 Factor2 Factor3
## trustNatGov      0.86
## trustParty       0.87
## trustParliament  0.85
## trustMilitary    0.71
## trustNewspaper           0.81
## trustTV                  0.85
## trustNeighbor                    0.64
## trustOther                       0.56
## trustCourt       0.44            0.32
## trustCivService
## trustPolice      0.48    0.30    0.38
## trustLocGov      0.44    0.32    0.37
## trustNGO                -0.49   -0.35
## trustRelative                    0.46
##
##                Factor1 Factor2 Factor3
## SS loadings      3.52    2.03    1.60
## Proportion Var   0.25    0.15    0.11
## Cumulative Var   0.25    0.40    0.51
##
## Test of the hypothesis that 3 factors are sufficient.
## The chi square statistic is 808 on 52 degrees of freedom.
## The p-value is 4e-136
```

1.6 构建二阶相关矩阵 (Square Multiple Correlation, SMC)，是否存在 SMC 过小的变量？观察反映象协方差矩阵 (Anti-Image Covariance Coefficients)，是否存在过多的太大的系数？观察 Kaiser–Meyer–Olkin 统计量，是否存在 KMO 过小 (KMO < 0.6) 的变量？如果以上统计量有不符合标准的，对变量进行调整（删除），重新进行因子分析。

```
smc(data)
```

```
##    trustCourt    trustNatGov      trustParty trustParliament trustCivService
```

```
##          0.392           0.691           0.703           0.684           0.095
##     trustMilitary      trustPolice       trustLocGov    trustNewspaper        trustTV
##          0.519           0.530           0.478           0.635           0.625
##        trustNGO     trustRelative     trustNeighbor        trustOther
##          0.320           0.195           0.351           0.260
```

```
cov2cor(solve(corMat))
```

```
##                   trustCourt trustNatGov trustParty trustParliament
## trustCourt           1.0000     -0.1055     0.0057        -0.03193
## trustNatGov         -0.1055      1.0000    -0.4450        -0.27659
## trustParty           0.0057     -0.4450     1.0000        -0.36332
## trustParliament     -0.0319     -0.2766    -0.3633         1.00000
## trustCivService     -0.0584     -0.0313     0.0175        -0.02267
## trustMilitary       -0.0213     -0.1408    -0.1308        -0.23402
## trustPolice         -0.2666     -0.0228    -0.0309        -0.03242
## trustLocGov         -0.1307      0.0099    -0.0238        -0.08363
## trustNewspaper      -0.0409     -0.0593     0.0205         0.02376
## trustTV             -0.0321      0.0098    -0.0347        -0.03752
## trustNGO             0.0314     -0.0805    -0.0097        -0.00069
## trustRelative       -0.0560     -0.0137    -0.0127        -0.01983
## trustNeighbor        0.0131     -0.0370    -0.0216         0.03360
## trustOther          -0.0172      0.0081     0.0212        -0.01689
##                   trustCivService trustMilitary trustPolice trustLocGov
## trustCourt                -0.0584      -0.02127     -0.2666     -0.1307
## trustNatGov               -0.0313      -0.14078     -0.0228      0.0099
## trustParty                 0.0175      -0.13082     -0.0309     -0.0238
## trustParliament           -0.0227      -0.23402     -0.0324     -0.0836
## trustCivService            1.0000      -0.02586     -0.0201     -0.0251
## trustMilitary             -0.0259       1.00000     -0.1437     -0.0499
## trustPolice               -0.0201      -0.14368      1.0000     -0.3587
## trustLocGov               -0.0251      -0.04989     -0.3587      1.0000
## trustNewspaper            -0.0084      -0.02009     -0.0348     -0.0218
## trustTV                   -0.0279      -0.00027     -0.0242     -0.0906
## trustNGO                   0.0713      -0.01584      0.0665      0.0509
## trustRelative             -0.0226      -0.02369     -0.0099     -0.0034
## trustNeighbor             -0.1060      -0.01789      0.0079     -0.1022
## trustOther                 0.0753       0.05387     -0.0907      0.0183
##                   trustNewspaper  trustTV trustNGO trustRelative trustNeighbor
```

```
## trustCourt              -0.0409 -0.03208  0.03145        -0.0560          0.0131
## trustNatGov             -0.0593  0.00978 -0.08045        -0.0137         -0.0370
## trustParty               0.0205 -0.03470 -0.00974        -0.0127         -0.0216
## trustParliament          0.0238 -0.03752 -0.00069        -0.0198          0.0336
## trustCivService         -0.0084 -0.02788  0.07131        -0.0226         -0.1060
## trustMilitary           -0.0201 -0.00027 -0.01584        -0.0237         -0.0179
## trustPolice             -0.0348 -0.02416  0.06647        -0.0099          0.0079
## trustLocGov             -0.0218 -0.09062  0.05094        -0.0034         -0.1022
## trustNewspaper           1.0000 -0.64783  0.18451        -0.0211         -0.0820
## trustTV                 -0.6478  1.00000  0.14092         0.0188          0.0149
## trustNGO                 0.1845  0.14092  1.00000         0.0274          0.0499
## trustRelative           -0.0211  0.01877  0.02744         1.0000         -0.2748
## trustNeighbor           -0.0820  0.01489  0.04994        -0.2748          1.0000
## trustOther              -0.0424  0.01943  0.15278        -0.0816         -0.3426
##                 trustOther
## trustCourt         -0.0172
## trustNatGov         0.0081
## trustParty          0.0212
## trustParliament    -0.0169
## trustCivService     0.0753
## trustMilitary       0.0539
## trustPolice        -0.0907
## trustLocGov         0.0183
## trustNewspaper     -0.0424
## trustTV             0.0194
## trustNGO            0.1528
## trustRelative      -0.0816
## trustNeighbor      -0.3426
## trustOther          1.0000
```

```
KMO(data)
```

```
## Kaiser-Meyer-Olkin factor adequacy
## Call: KMO(r = data)
## Overall MSA =  0.88
## MSA for each item =
##      trustCourt      trustNatGov      trustParty trustParliament trustCivService
##            0.94             0.89            0.88            0.90            0.93
##   trustMilitary      trustPolice      trustLocGov  trustNewspaper         trustTV
```

```
##             0.95            0.91            0.92            0.81            0.81
##        trustNGO    trustRelative   trustNeighbor      trustOther
##             0.91            0.88            0.83            0.80
```

```
# cov2cor(KMO(data)$ImCov)
```

可以发现，smc 分析中 trustCivService、trustRelative 的值过小。KMO 分析中没有小于 0.8 的值。反像
相关矩阵检验检验中，trustNatGov、trustParty 两项的值太大，不适合做因子分析。

```
fac_data <- data[, -c(5, 12, 2, 3)]
M1 <- factanal(~., data = fac_data, factors = 3 , scores = "regression")
M2 <- update(M1, factor = 3, roation = "varimax", SMC = T)
print(M2, digits=2, cutoff=.3, sort=TRUE)
```

```
##
## Call:
## factanal(x = ~., factors = 3, data = fac_data, scores = "regression",    factor = 3, roation =
##
## Uniquenesses:
##      trustCourt trustParliament     trustMilitary      trustPolice     trustLocGov
##            0.59            0.40            0.41            0.43            0.49
##  trustNewspaper          trustTV         trustNGO   trustNeighbor      trustOther
##            0.23            0.22            0.63            0.61            0.57
##
## Loadings:
##                 Factor1 Factor2 Factor3
## trustCourt       0.55
## trustParliament  0.76
## trustMilitary    0.76
## trustPolice      0.65            0.30
## trustLocGov      0.60
## trustNewspaper           0.81
## trustTV                  0.83
## trustNeighbor                    0.57
## trustOther                       0.64
## trustNGO                 -0.47   -0.36
##
##                 Factor1 Factor2 Factor3
## SS loadings        2.43    1.82    1.16
```

```
## Proportion Var     0.24     0.18     0.12
## Cumulative Var     0.24     0.43     0.54
##
## Test of the hypothesis that 3 factors are sufficient.
## The chi square statistic is 409 on 18 degrees of freedom.
## The p-value is 1.3e-75
```
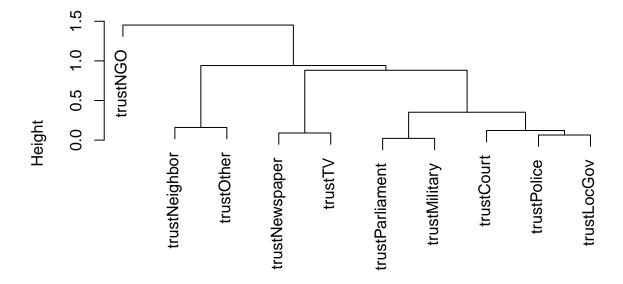
1.7 观察结构矩阵，对得出的因子进行命名。

PC1 为政府机构信任，PC2 为媒体信任，PC3 为个体信任。

1.8 构建因子得分变量 (变量名称使用上述命名)，并使用树状图进行可视化描述。

```
hc = hclust(dist(M2$loadings))
plot(hc)
```

Cluster Dendrogram

dist(M2$loadings)
hclust (*, "complete")