Homework 5

陈旭鹏

2018/5/20

1

$$\text{Since} \mathbf{Z} \text{is the standardized r.v.s, and } L = [.9, .7, .5]^T,$$
$$\text{we have the following:}$$

$$\mathbf{LL^T + \Psi} = [.9, .7, .5]^T[.9, .7, .5] + \begin{bmatrix} .19 & 0 & 0 \\ 0 & .51 & 0 \\ 0 & 0 & .75 \end{bmatrix}$$

$$= \begin{bmatrix} .81 & .63 & .45 \\ .63 & .49 & .35 \\ .45 & .35 & .25 \end{bmatrix} \begin{bmatrix} 1.0 & .63 & .45 \\ .63 & 1.0 & .35 \\ .45 & .35 & 1.0 \end{bmatrix} = \rho$$

2

a

Since m =1, the communalities are calculated as the following:

$$h_1^2 = \ell_{11}^2 = .9 = .81$$
$$h_2^2 = \ell_{21}^2 = .7 = .49$$
$$h_3^2 = \ell_{31}^2 = .5 = .25$$

Communalities indicate the common variance shared by factors with given variables. Higher communality indicated that larger amount of the variance in the variable has been extracted by the factor solution. For better measurement of factor analysis communalities should be 0.4 or greater.

b

$$Corr(Z_i, F_1) = \ell_{i1} \implies Corr(\mathbf{Z_i}, F_1) = \mathbf{L} = [.9, .7, .5]^T$$

Because the first variable Z_1 has the largest correlation with common factor, Z_1 will carry greatest weight in term of F_1.

3

In PCA: $X_{n\times p} \rightarrow \hat{\Sigma} = XX^T$

Applying spectral decompostion, we have: $\hat{\Sigma} = P\Lambda P^T$

PC scores form the matrix which is: $P\Lambda^{\frac{1}{2}}$

In MDS: $X_{n\times p} \rightarrow \mathbf{B} \rightarrow \mathbf{D}$

$$B = XX^T = P\Lambda P^T \rightarrow X = P\Lambda^{\frac{1}{2}}$$

So they are equivalent

4

```
library(MASS)
data <-source("table5_12.txt")$value
attach(data)
data
```

```
##                         1assault and battery 2rape 3embezzlement 4perjury
## 1assault and battery                     0.0  21.0          71.2     36.4
## 2rape                                   21.0   0.0          54.1     36.4
## 3embezzlement                           71.2  54.1           0.0     36.4
## 4perjury                                36.4  36.4          36.4      0.0
## 5libel                                  52.1  54.1          52.1      0.7
## 6burglary                               89.9  75.2          36.4     54.1
## 7prostitution                           53.0  73.0          75.2     52.1
## 8receiving stolen goods                 90.1  93.2          71.2     63.4
##                         5libel 6burglary 7prostitution
## 1assault and battery      52.1      89.9          53.0
## 2rape                     54.1      75.2          73.0
## 3embezzlement             52.1      36.4          75.2
## 4perjury                   0.7      54.1          52.1
## 5libel                     0.0      53.0          36.4
## 6burglary                 53.0       0.0          88.0
## 7prostitution             36.4      88.0           3.0
## 8receiving stolen goods   52.1      36.4          73.0
##                         8receiving stolen goods
## 1assault and battery                       90.1
## 2rape                                      93.2
## 3embezzlement                              71.2
## 4perjury                                   63.4
## 5libel                                     52.1
## 6burglary                                  36.4
## 7prostitution                              73.0
## 8receiving stolen goods                     0.0
```

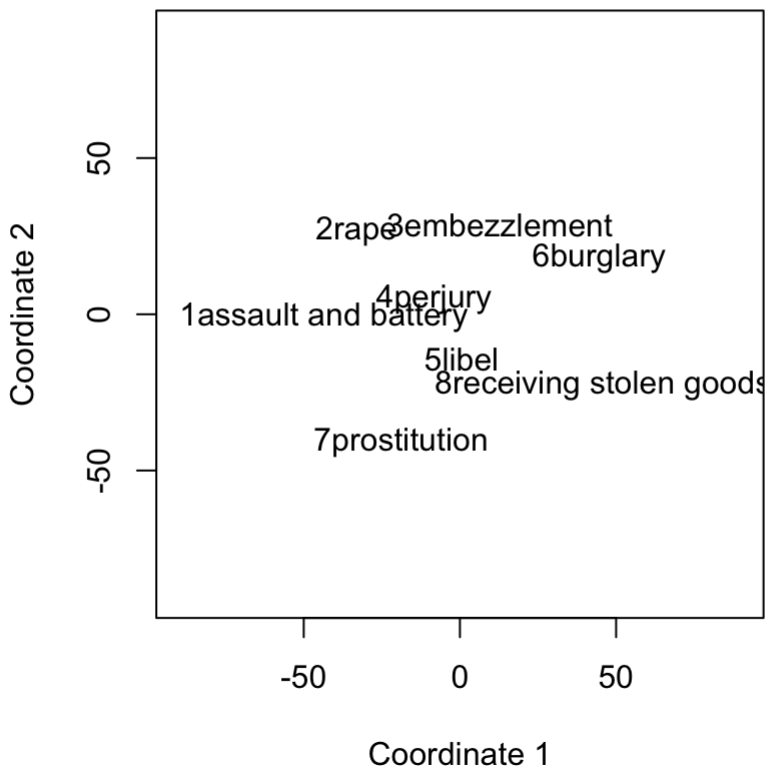
```
data.mds<-cmdscale(as.matrix(data),k=2,eig=T)
data.mds$eig
```

```
## [1]  7.656184e+03  4.284887e+03  1.247778e+03  4.551195e+02  4.458949e+01
## [6]  6.252776e-13 -4.294064e+02 -4.625762e+02
```

```
data.mds$points
```

```
##                               [,1]        [,2]
## 1assault and battery     43.5559886  -0.7119682
## 2rape                    33.0203254  26.8701326
## 3embezzlement           -12.7175678  28.6671549
## 4perjury                  8.2385788   4.9823016
## 5libel                   -0.5537496 -14.3945452
## 6burglary               -44.5558031  18.0592801
## 7prostitution            18.8280191 -40.9023243
## 8receiving stolen goods -45.8157912 -22.5700315
```

```
par(pty="s")
xyrange = 90
plot(-data.mds$points[,1],data.mds$points[,2],type="n",xlab="Coordinate 1",ylab="Coor
dinate 2",
xlim=c(-xyrange,xyrange),ylim=c(-xyrange,xyrange))
text(-data.mds$points[,1],data.mds$points[,2],labels=row.names(data))
```



The two dimensions can be interpreted as the degree of immorality and the severity of the consequences.