Homework 4

Zahlen Zbinden

2023-11-28

1. A study of aged 65 and older was conducted on a sample of generally healthy adults, randomly selected from Medicare rolls. A number of variables were measured for each subject.

a. Perform a principal components factor analysis based on the given correlation matrix, for m = 2 and m = 3 factors. Describe how you might interpret the resulting factors for each model: can you describe the underlying latent variables for these two models? Which variables contribute most to each factor?

```
physio <- read.table("D:/RepoMan/osu/data/PhysioData.csv", sep = ",")
```

first lets look for interrelated pair in the corelation matrix:

We can see there is a cluster at fev with height, and weight. There is also less correletated cluster at crt with height and weight. There is another cluster with dsst, fev, and aai. We could say the correlated sets of variables at crt, weight, and height as well as dsst, fev, and aai.

```
round(physio, 2)
```

	weight	height	physact	ldl	alb	crt	plt	sbp	aai	fev	dsst
weight	1.00	0.55	-0.03	0.00	0.05	0.25	-0.15	0.01	0.09	0.34	0.06
height	0.55	1.00	0.06	-0.16	0.09	0.36	-0.29	-0.07	0.07	0.58	0.02
physact	-0.03	0.06	1.00	-0.03	0.01	-0.03	-0.01	0.00	0.08	0.10	-0.03
ldl	0.00	-0.16	-0.03	1.00	0.12	-0.13	0.20	-0.03	-0.06	-0.06	0.00
alb	0.05	0.09	0.01	0.12	1.00	0.04	-0.06	-0.02	0.03	0.07	0.07
crt	0.25	0.36	-0.03	-0.13	0.04	1.00	-0.15	0.00	-0.05	0.22	-0.15
plt	-0.15	-0.29	-0.01	0.20	-0.06	-0.15	1.00	0.03	-0.08	-0.18	0.02
sbp	0.01	-0.07	0.00	-0.03	-0.02	0.00	0.03	1.00	-0.33	-0.11	-0.16
aai	0.09	0.07	0.08	-0.06	0.03	-0.05	-0.08	-0.33	1.00	0.23	0.21
fev	0.34	0.58	0.10	-0.06	0.07	0.22	-0.18	-0.11	0.23	1.00	0.16

```
dsst      0.06    0.02    -0.03   0.00    0.07   -0.15   0.02  -0.16   0.21   0.16   1.00
atrophy   0.06    0.12    -0.08  -0.02    0.05    0.15  -0.06   0.06  -0.08  -0.01   0.01
          atrophy
weight      0.06
height      0.12
physact    -0.08
ldl        -0.02
alb         0.05
crt         0.15
plt        -0.06
sbp         0.06
aai        -0.08
fev        -0.01
dsst        0.01
atrophy     1.00
```

```
physio |>
  ggcorrplot()
```
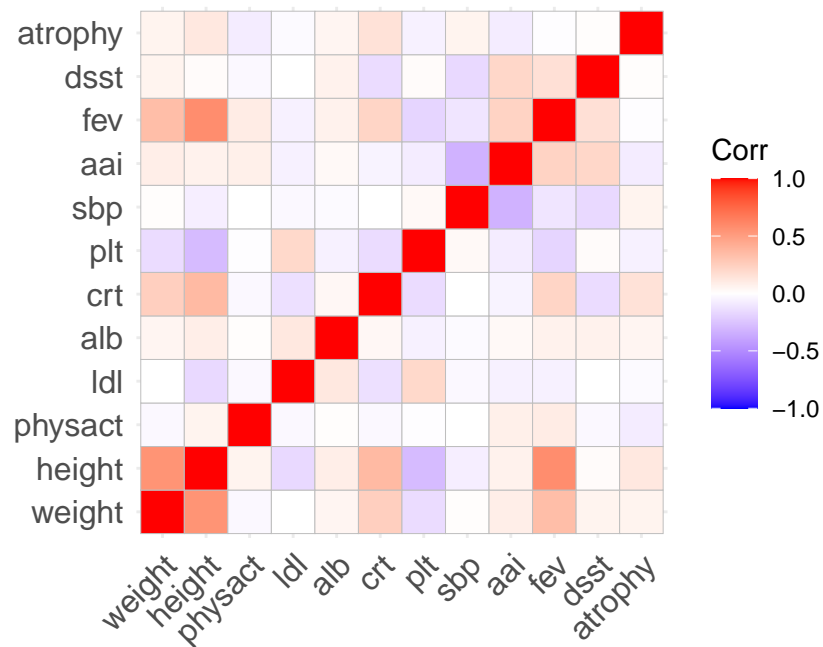


```
pca_2 <- principal(r = physio, nfactors = 2, scores = TRUE)
pca_2
```

```
Principal Components Analysis
Call: principal(r = physio, nfactors = 2, scores = TRUE)
Standardized loadings (pattern matrix) based upon correlation matrix
          RC1   RC2    h2   u2 com
weight   0.67  0.09 0.462 0.54 1.0
height   0.85  0.14 0.749 0.25 1.1
physact  0.03  0.16 0.028 0.97 1.1
ldl     -0.26  0.05 0.071 0.93 1.1
alb      0.12  0.12 0.027 0.97 2.0
crt      0.62 -0.23 0.439 0.56 1.3
plt     -0.47 -0.04 0.221 0.78 1.0
sbp      0.02 -0.64 0.411 0.59 1.0
aai      0.06  0.74 0.544 0.46 1.0
fev      0.64  0.40 0.574 0.43 1.7
dsst    -0.05  0.58 0.337 0.66 1.0
atrophy  0.25 -0.28 0.141 0.86 2.0


                        RC1   RC2
SS loadings            2.35 1.65
Proportion Var         0.20 0.14
Cumulative Var         0.20 0.33
Proportion Explained   0.59 0.41
Cumulative Proportion 0.59 1.00


Mean item complexity =  1.3
Test of the hypothesis that 2 components are sufficient.


The root mean square of the residuals (RMSR) is  0.09


Fit based upon off diagonal values = 0.68
```

```r
pca_2_rotate <- principal(r = physio, nfactors = 2, rotate = "none", scores = TRUE)
pca_2_rotate
```

```
Principal Components Analysis
Call: principal(r = physio, nfactors = 2, rotate = "none", scores = TRUE)
Standardized loadings (pattern matrix) based upon correlation matrix
          PC1   PC2    h2   u2 com
weight   0.67 -0.11 0.462 0.54 1.1
height   0.86 -0.13 0.749 0.25 1.0
physact  0.08  0.14 0.028 0.97 1.6
```

```
ldl     -0.23  0.13 0.071 0.93 1.5
alb      0.14  0.08 0.027 0.97 1.5
crt      0.52 -0.41 0.439 0.56 1.9
plt     -0.46  0.11 0.221 0.78 1.1
sbp     -0.18 -0.62 0.411 0.59 1.2
aai      0.28  0.68 0.544 0.46 1.3
fev      0.73  0.19 0.574 0.43 1.1
dsst     0.13  0.57 0.337 0.66 1.1
atrophy  0.15 -0.34 0.141 0.86 1.4


                        PC1  PC2
SS loadings            2.43 1.57
Proportion Var         0.20 0.13
Cumulative Var         0.20 0.33
Proportion Explained   0.61 0.39
Cumulative Proportion  0.61 1.00

Mean item complexity =  1.3
Test of the hypothesis that 2 components are sufficient.

The root mean square of the residuals (RMSR) is  0.09

Fit based upon off diagonal values = 0.68
```