Package *MFA01* Presentation

Sheng Liu, Shiying Zhou, Wenzhe Ding, Haochi Zhang

12/1/2016

About the presentation

In this presentation, we will briefly introduce the following things

- summary of the package tools
- a brief introduction of multiple factor analysis
- showcase of the package functionalities

About MFA01 package

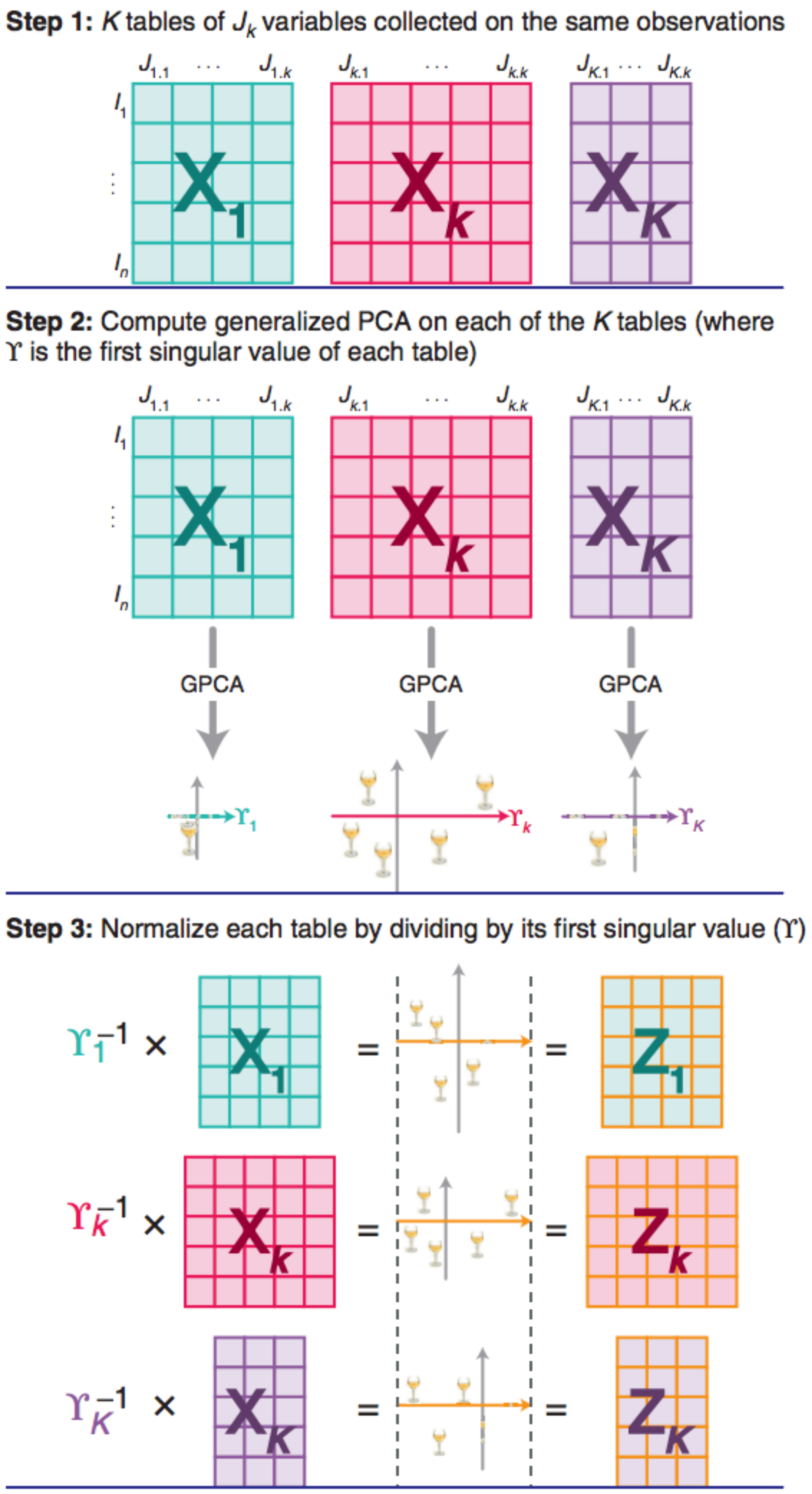
In this package, we created tools for multiple factor analysis (MFA), an extension of principle component analysis (PCA) to process multiple data sets and their correlations. The package include tools to compute the following outputs from MFA

- eigenvalues, common factor scores, partial factor scores, loadings
- visualization tools: generic functions for print, plot and tabularizing results
- auxiliary functions: *summaries of eigenvalues*, *contributions*, R_V *coefficients*, *coefficients to study the between table structure*, L_q *coefficients* and *Bootstrap* method to estimate the stability of the compromise factor scores.
- A shiny application

What is MFA?

- Multiple factor analysis is a generalization of principle component analysis.
- The goal of MFA is to “analyze several data sets of variables collected on the same set of observations, orâ”as in its dual versionâseveral sets of observations measured on the same set of variables"
- MFA can be summarized into five steps, nicely illustrated by figure 1 by Abdi et al. (2013):
- Step 1: K tables of J_k variables collected on the same observations
- Step 2: Compute generalized PCA on each of the K tables (where Ï is the first singular value of each table)
- Step 3: Normalize each table by dividing by its first singular value (Ï)
- Step 4: Concatenate the K normalized tables
- Step 5: Compute a generalized PCA on the concatenated table

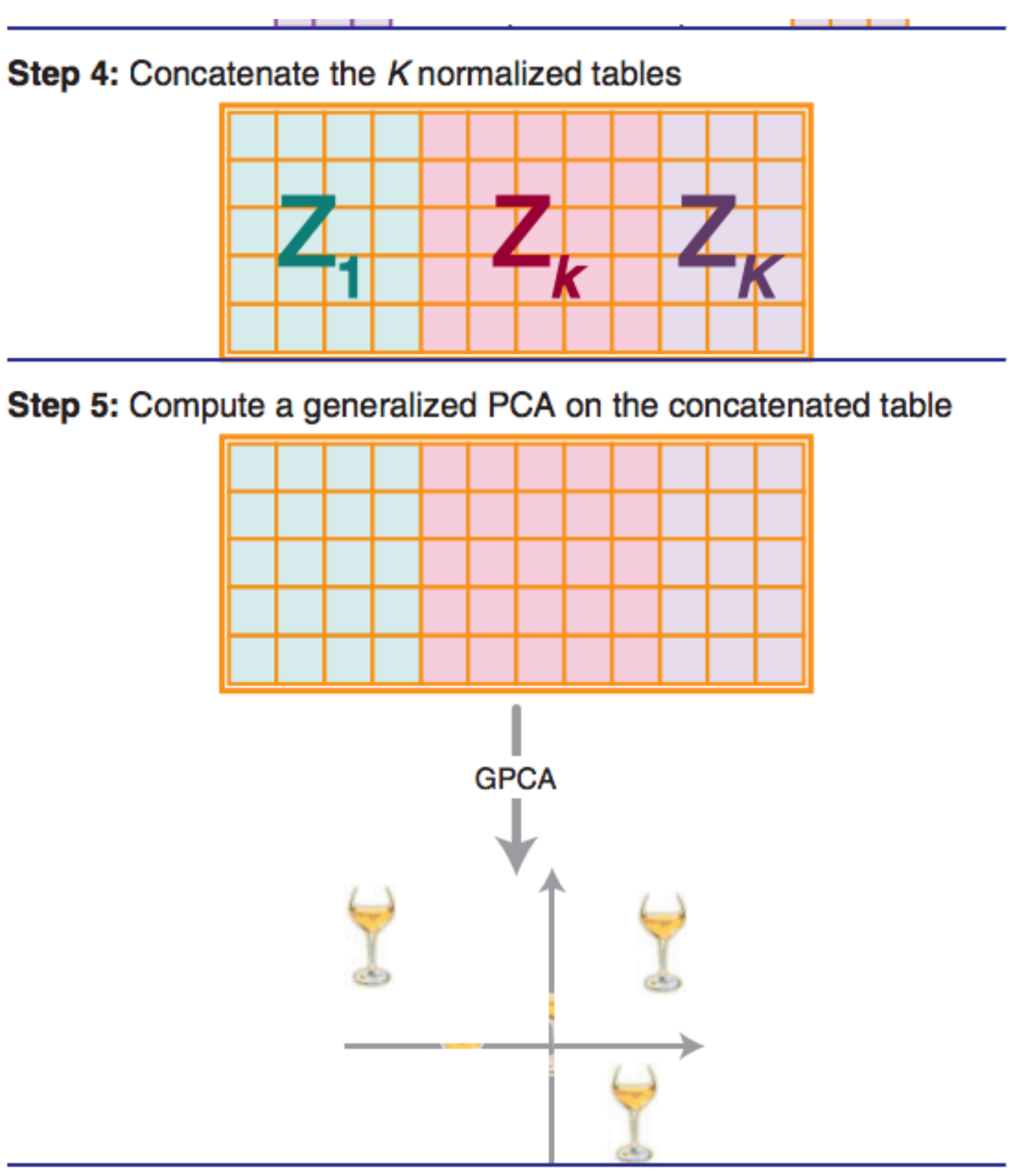
What is MFA?



Step 1: K tables of J_k variables collected on the same observations
$J_{1.1}$ … $J_{1.k}$ $J_{k.1}$ … $J_{k.k}$ $J_{K.1}$ … $J_{K.k}$
I_1
I_n
X_1 X_k X_K
Step 2: Compute generalized PCA on each of the K tables (where Υ is the first singular value of each table)
GPCA
GPCA
GPCA
Υ_1 Υ_k Υ_K
Step 3: Normalize each table by dividing by its first singular value (Υ)
$\Upsilon_1^{-1} \times X_1 = Z_1$
$\Upsilon_k^{-1} \times X_k = Z_k$
$\Upsilon_K^{-1} \times X_K = Z_K$

step 1 to 3

What is MFA?


Step 4: Concatenate the *K* normalized tables
Z_1
Z_k
Z_K
Step 5: Compute a generalized PCA on the concatenated table
GPCA

step 4 to 5

How to install the package?

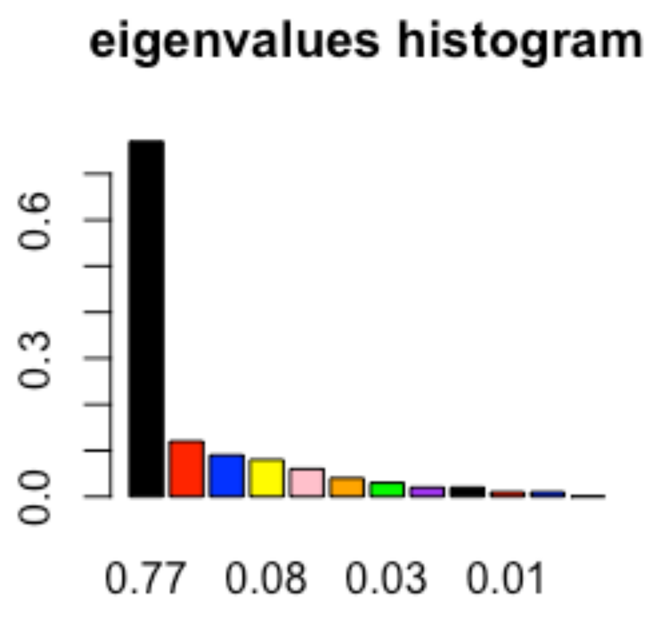
- To install the package, go to the R file ‘install.R’, use ‘setwd()’ to change the working directory to ‘your_path/MFA01’, and run the file.
- Then the MFA package should be installed. To create an ‘mfa’ object with default data set, simply run

```
## object "mfa"
##
## Number of tables/blocks for analysis: 10
## Weight of tables:
##  0.240983 0.2386034 0.2748748 0.2728329 0.3065003 0.3023908 0.4167984 0.2724079 0.2635522 0.308608
## First two eigenvalues:  0.7702551 0.1229254
## First component of common factor scores:
##  -0.9802057 -0.8088651 -0.7610058 -1.114984 1.372757 1.264015 0.8082827 0.9253423 -0.6689538 0.07316059 -0.4761
## Facor loadings for the first table (1st component):
##  -0.2943913 -0.2665541 -0.2599726 0.2411168 0.2863237 -0.232907
```

Functionalities

The mfa object computes the eigenvalues of the grand data table, which can be visualized by

```
plot_eig(mfa_obj)
```



Functionalities

A matrix of the common factor score can be accessed by

```
mfa_obj$cfs
```

```
##               [,1]        [,2]        [,3]         [,4]         [,5]
##  [1,] -0.98020575  0.16325474 -0.02833247 -0.122967423 -0.138960264
##  [2,] -0.80886515  0.03262348  0.16181752 -0.376869178 -0.201380272
##  [3,] -0.76100584 -0.45418702 -0.00567849 -0.134631905 -0.061412658
##  [4,] -1.11498367 -0.16586214 -0.22362954  0.567155437  0.006220427
##  [5,]  1.37275684 -0.12838880  0.12163879 -0.315803256 -0.228652248
##  [6,]  1.26401538 -0.10813651 -0.37010578  0.087703704  0.413438110
##  [7,]  0.80828274  0.20466790  0.25317711 -0.085518365  0.247645688
##  [8,]  0.92534231  0.40775212 -0.37972251  0.264344947 -0.505067639
##  [9,] -0.66895382  0.36852275 -0.43311223 -0.342059710  0.247480544
## [10,]  0.07316059 -0.75677932 -0.04550557 -0.003904788  0.043136934
## [11,] -0.47610885  0.51276640  0.40040815  0.173699464  0.217286089
## [12,]  0.36656519 -0.07623359  0.54904503  0.288851074 -0.039734711
##               [,6]         [,7]        [,8]        [,9]        [,10]
##  [1,]  0.21752965 -0.220051492  0.26947251 -0.21850438  0.077474226
##  [2,] -0.13572321  0.147904830 -0.02638970  0.22079673  0.123996699
##  [3,] -0.07115109 -0.279400608 -0.03095015  0.15088378 -0.008565486
##  [4,] -0.30152727  0.002219817  0.01083609 -0.03067018 -0.117655827
##  [5,] -0.06585165 -0.138077283 -0.07701356 -0.08147074 -0.230493449
##  [6,]  0.12748875 -0.156705743  0.08745239  0.16910491  0.072655724
##  [7,] -0.42605557  0.046374939  0.09309193 -0.14630764  0.122837767
##  [8,]  0.03296264  0.058573490 -0.07536367  0.02332136  0.095377250
##  [9,]  0.07862601  0.250187914  0.01424231 -0.00850599 -0.155918796
## [10,]  0.17500807  0.223444581 -0.18009379 -0.18214255  0.104045535
## [11,]  0.16511358 -0.138517707 -0.32629392 -0.01128101  0.006704255
## [12,]  0.20358008  0.204047262  0.24100955  0.11477572 -0.090457898
##               [,11]        [,12]
##  [1,] -0.059739737 1.580398e-16
##  [2,] -0.162685833 1.580398e-16
##  [3,]  0.221691315 1.580398e-16
##  [4,] -0.104911814 1.580398e-16
##  [5,] -0.094639543 1.580398e-16
##  [6,] -0.101440700 1.580398e-16
##  [7,]  0.086496068 1.580398e-16
##  [8,]  0.087305272 1.580398e-16
##  [9,]  0.087365252 1.580398e-16
## [10,] -0.001839931 1.580398e-16
## [11,] -0.014311737 1.580398e-16
## [12,]  0.056711388 1.580398e-16
```

Functionalities

Other elements of the mfa object include:

```
attributes(mfa_obj)$names
```

```
## [1] "assessors"   "index_lists" "weights"     "eigen"       "cfs"
## [6] "fl"          "pfl"         "pfs"
```

Functionalities

To see a summary of what the object returns

```
mfa_obj
```

```
## object "mfa"
##
## Number of tables/blocks for analysis: 10
## Weight of tables:
##  0.240983 0.2386034 0.2748748 0.2728329 0.3065003 0.3023908 0.4167984 0.2724079 0.2635522 0.308608
## First two eigenvalues:  0.7702551 0.1229254
## First component of common factor scores:
##  -0.9802057 -0.8088651 -0.7610058 -1.114984 1.372757 1.264015 0.8082827 0.9253423 -0.6689538 0.07316059 -0.4761
## Facor loadings for the first table (1st component):
##  -0.2943913 -0.2665541 -0.2599726 0.2411168 0.2863237 -0.232907
```

Eigenvalues

Our “ev.summary”" takes the “mfa” object and returns a table with the singular values, the eigenvalues, cumulative, percentage of intertia, cumulative percentage of inertia, for all the extracted components.

```r
setwd("C:/Users/Sheng/Google Drive/Berkeley16Fall/STAT243/Stats-243/Stats-243/final")
library(MFA01)
mfa_obj <- MFA()
data <- get(load("data/wine.rda"))
ev.summary(mfa_obj)
```

```
##    SingularValue Eigenvalue CumulativeEigenvalue Inertia CumulativeInertia
## 1           0.88       0.77                 0.77      62                62
## 2           0.35       0.12                 0.89      10                72
## 3           0.30       0.09                 0.98       7                79
## 4           0.28       0.08                 1.06       6                85
## 5           0.24       0.06                 1.12       5                90
## 6           0.20       0.04                 1.16       3                93
## 7           0.17       0.03                 1.19       2                95
## 8           0.14       0.02                 1.21       2                97
## 9           0.14       0.02                 1.23       2                99
## 10          0.10       0.01                 1.24       1               100
## 11          0.10       0.01                 1.25       1               101
## 12          0.00       0.00                 1.25       0               101
```

Contributions

We include three functions to calculate *contributions*: 1. "ctr.obs": contribution of an observation to a dimension 1. "ctr.var": contribution of a variable to a dimension 1. "ctr.table": contribution of a table to a dimension

For instance, we call "ctr.obs" as follows:

```
ctr.obs(mfa_obj)
```

```
##            [,1]     [,2]     [,3]     [,4]     [,5]    [,6]     [,7]
##  [1,] 0.103949 0.018068 7.37e-04 1.66e-02 2.70e-02 0.10059 1.31e-01
##  [2,] 0.070784 0.000722 2.41e-02 1.56e-01 5.67e-02 0.03916 5.90e-02
##  [3,] 0.062656 0.139845 2.96e-05 1.99e-02 5.27e-03 0.01076 2.10e-01
##  [4,] 0.134500 0.018650 4.59e-02 3.53e-01 5.41e-05 0.19326 1.33e-05
##  [5,] 0.203878 0.011175 1.36e-02 1.09e-01 7.31e-02 0.00922 5.14e-02
##  [6,] 0.172858 0.007927 1.26e-01 8.43e-03 2.39e-01 0.03455 6.62e-02
##  [7,] 0.070682 0.028397 5.89e-02 8.02e-03 8.57e-02 0.38586 5.80e-03
##  [8,] 0.092638 0.112712 1.32e-01 7.66e-02 3.57e-01 0.00231 9.25e-03
##  [9,] 0.048415 0.092067 1.72e-01 1.28e-01 8.56e-02 0.01314 1.69e-01
## [10,] 0.000579 0.388254 1.90e-03 1.67e-05 2.60e-03 0.06510 1.35e-01
## [11,] 0.024524 0.178244 1.47e-01 3.31e-02 6.60e-02 0.05795 5.17e-02
## [12,] 0.014537 0.003940 2.77e-01 9.15e-02 2.21e-03 0.08810 1.12e-01
##            [,8]     [,9]    [,10]    [,11]  [,12]
##  [1,] 0.242454 0.213206 0.037223 0.026319 0.0833
##  [2,] 0.002325 0.217703 0.095350 0.195187 0.0833
##  [3,] 0.003198 0.101663 0.000455 0.362450 0.0833
##  [4,] 0.000392 0.004201 0.085848 0.081171 0.0833
##  [5,] 0.019803 0.029640 0.329472 0.066054 0.0833
##  [6,] 0.025535 0.127700 0.032737 0.075888 0.0833
##  [7,] 0.028935 0.095590 0.093576 0.055175 0.0833
##  [8,] 0.018964 0.002429 0.056415 0.056212 0.0833
##  [9,] 0.000677 0.000323 0.150764 0.056290 0.0833
## [10,] 0.108292 0.148150 0.067135 0.000025 0.0833
## [11,] 0.355483 0.000568 0.000279 0.001511 0.0833
## [12,] 0.193941 0.058827 0.050745 0.023719 0.0833
```

Coefficients to study the Between-Table Structure

To evaluate the similarity between two tables we can use the R_V coefficient.

```r
table1=mfa_obj$assessors[[1]]
table2=mfa_obj$assessors[[2]]
RV(table1,table2)
```

```
## [1] 0.868
```

Coefficients to study the Between-Table Structure (Continued)

We can also get a matrix of R_V coefficients by the following function.

```
RV_table(data, sets = list(2:3, 4:5, 6:10))
```

```
##        [,1]  [,2]  [,3]
## [1,] 1.000 0.907 0.984
## [2,] 0.907 1.000 0.946
## [3,] 0.984 0.946 1.000
```

L_g Coefficients

We design two functions for L_g coefficients similar to R_V coefficients.

```
Lg(table1,table2)
```

```
## [1] 0.918
```

```
Lg_table(data, sets = list(2:3, 4:5, 6:10))
```

```
##       [,1]  [,2]  [,3]
## [1,] 1.000 0.919 0.988
## [2,] 0.919 1.027 0.962
## [3,] 0.988 0.962 1.006
```

Bootstrap

We also write a function that allows the user to perform bootstrapping in order to estimate the stability of the compromise factor scores.

```
bt <- bootStrap(mfa_obj)
bt$mean
```

```
##          [,1]    [,2]     [,3]    [,4]    [,5]     [,6]     [,7]    [,8]
##  [1,] -0.835  0.1148 -0.01182 -0.0855 -0.1263  0.19103 -0.17988  0.2323
##  [2,] -0.689  0.0233  0.16444 -0.3114 -0.1796 -0.11426  0.14213 -0.0185
##  [3,] -0.640 -0.3790  0.00248 -0.1284 -0.0598 -0.03339 -0.24867 -0.0163
##  [4,] -0.939 -0.0910 -0.19046  0.4524 -0.0243 -0.23342  0.00371  0.0153
##  [5,]  1.160 -0.1326  0.09211 -0.2605 -0.1667 -0.08168 -0.12563 -0.0577
##  [6,]  1.083 -0.0982 -0.34679  0.0563  0.3796  0.10013 -0.15675  0.0801
##  [7,]  0.692  0.1789  0.19104 -0.0778  0.2073 -0.38573  0.04137  0.0666
##  [8,]  0.778  0.3321 -0.30851  0.2508 -0.3887 -0.00272  0.06088 -0.0515
##  [9,] -0.563  0.2842 -0.37184 -0.2613  0.2187  0.04909  0.22861  0.0190
## [10,]  0.064 -0.6248 -0.05258 -0.0463  0.0336  0.17601  0.17982 -0.1684
## [11,] -0.408  0.4290  0.36337  0.1748  0.1650  0.14937 -0.09954 -0.2813
## [12,]  0.298 -0.0367  0.46856  0.2370 -0.0586  0.18557  0.15395  0.1804
##           [,9]    [,10]    [,11]     [,12]
##  [1,] -0.1769  0.05332 -0.05871 -0.014857
##  [2,]  0.1873  0.10948 -0.14849 -0.029820
##  [3,]  0.1285 -0.01183  0.19682 -0.008375
##  [4,] -0.0394 -0.12090 -0.08361  0.004921
##  [5,] -0.0746 -0.18047 -0.08662 -0.000736
##  [6,]  0.1405  0.06434 -0.07837  0.025160
##  [7,] -0.1153  0.10676  0.06853  0.004905
##  [8,]  0.0138  0.07102  0.07893 -0.007235
##  [9,] -0.0152 -0.12631  0.06961 -0.025214
## [10,] -0.1750  0.09433 -0.00200  0.019552
## [11,]  0.0102  0.00998 -0.00696  0.014939
## [12,]  0.1162 -0.06972  0.05086  0.016759
```

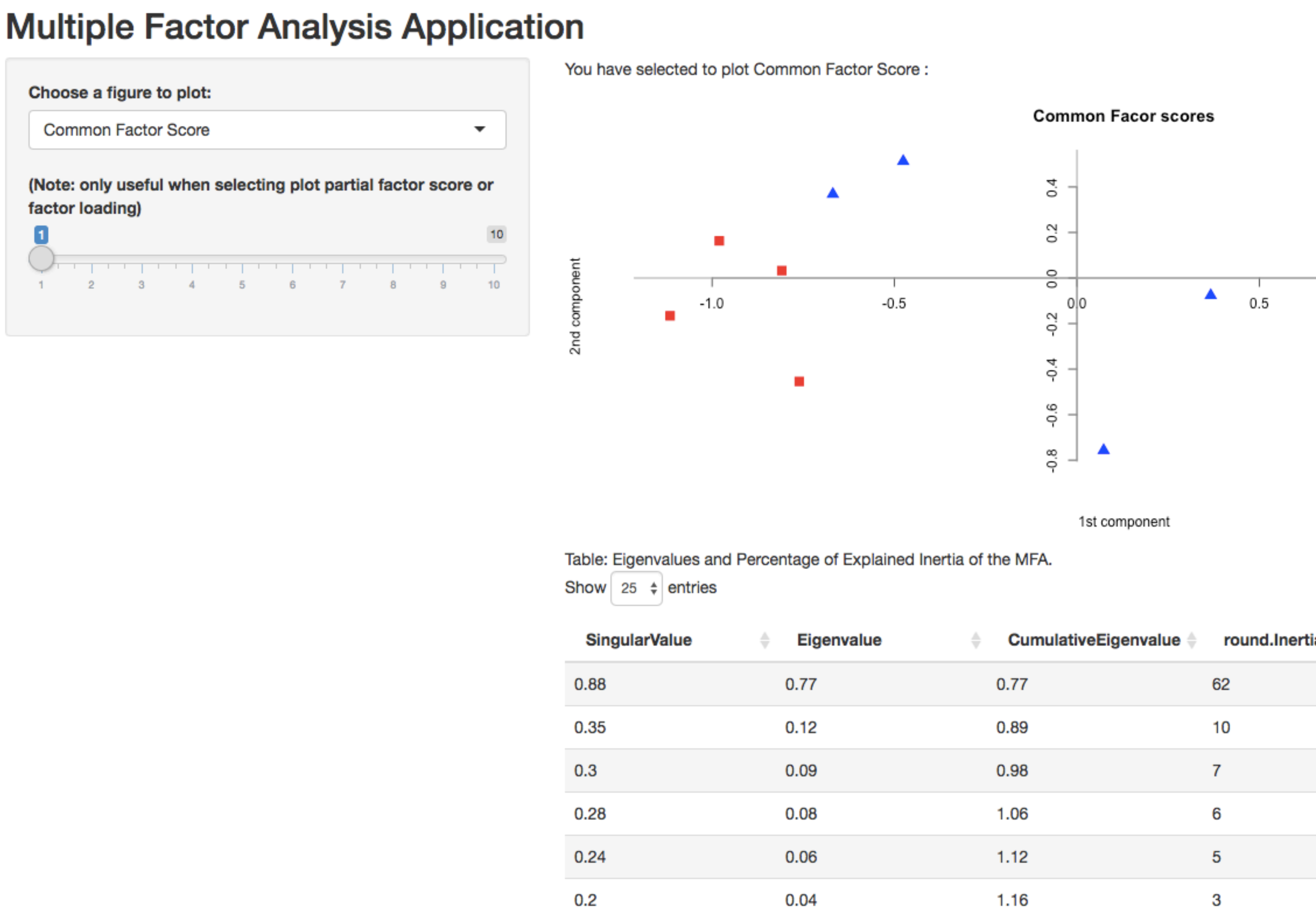
Bootstrap (Continued)

```
bt$sd
```

```
##         [,1]   [,2]   [,3]   [,4]   [,5]   [,6]   [,7]   [,8]   [,9]
##  [1,] 0.0526 0.0944 0.0571 0.1331 0.0774 0.1169 0.1432 0.1462 0.1488
##  [2,] 0.0841 0.0900 0.0904 0.1397 0.0941 0.1143 0.1237 0.1307 0.0901
##  [3,] 0.0810 0.0732 0.0801 0.1266 0.0854 0.1459 0.1346 0.0967 0.1217
##  [4,] 0.0684 0.1158 0.1045 0.1218 0.1355 0.2037 0.1238 0.1553 0.1166
##  [5,] 0.0737 0.1033 0.0791 0.1251 0.1446 0.2304 0.1337 0.1574 0.1861
##  [6,] 0.0872 0.1011 0.1552 0.2122 0.1490 0.1804 0.2013 0.1980 0.1649
##  [7,] 0.0915 0.0812 0.0752 0.0787 0.0916 0.1411 0.1111 0.1006 0.1342
##  [8,] 0.0559 0.0853 0.1284 0.1621 0.1433 0.1680 0.1613 0.1636 0.1376
##  [9,] 0.0881 0.1004 0.0720 0.1051 0.0678 0.1281 0.0905 0.0601 0.1156
## [10,] 0.0631 0.0990 0.0704 0.0877 0.0922 0.1045 0.0544 0.1025 0.1157
## [11,] 0.0720 0.0640 0.1154 0.1056 0.1116 0.1407 0.1177 0.0946 0.1475
## [12,] 0.0642 0.0983 0.0718 0.0654 0.0688 0.0945 0.0773 0.1067 0.0990
##        [,10]  [,11]  [,12]
##  [1,] 0.1392 0.2147 0.0709
##  [2,] 0.0849 0.2109 0.1045
##  [3,] 0.1058 0.1505 0.0846
##  [4,] 0.1702 0.1826 0.1326
##  [5,] 0.1668 0.2374 0.1031
##  [6,] 0.1461 0.2396 0.1072
##  [7,] 0.1131 0.1904 0.0651
##  [8,] 0.0967 0.1645 0.1080
##  [9,] 0.0824 0.1609 0.1130
## [10,] 0.0984 0.0763 0.1396
## [11,] 0.0826 0.0675 0.0877
## [12,] 0.0630 0.0783 0.0976
```

Shiny App

The interface of the app looks like



Multiple Factor Analysis Application

Choose a figure to plot: Common Factor Score

(Note: only useful when selecting plot partial factor score or factor loading)

1 — 10

You have selected to plot Common Factor Score :

Table: Eigenvalues and Percentage of Explained Inertia of the MFA.

Show 25 entries

SingularValue	Eigenvalue	CumulativeEigenvalue	round.Inerti
0.88	0.77	0.77	62
0.35	0.12	0.89	10
0.3	0.09	0.98	7
0.28	0.08	1.06	6
0.24	0.06	1.12	5
0.2	0.04	1.16	3

shiny_app_interface

One can use the drop down list to choose which plot to show and slide the number bar to choose which table to plot the partial factor scores and laoding.