**SL-VII ASSIGNMENT NO. 8**

Title: Dimensionality Reduction

Problem Statement:ImplementPrincipal Component Analysis-Finding Principal Components, Variance and Standard Deviation calculations of principal components.

Code and Outputs:

> #Data set Business Industry categories and their Progress

> #The database is attached to the R search path. This means that

> #the database is searched by R when evaluating a variable, so

> #objects in the database can be accessed by simply giving their

> #names.

> mydata<-read.csv("pca\_gsp.csv")

> attach(mydata)

The following objects are masked from mydata (pos = 3):

Ag, Comm, Constr, Energy, Govt, Manuf, Manuf\_nd,

Mining, RE, Services, State, TradeR, TradeW,

Transp

> # list the variables in mydata

> names(mydata)

[1] "State" "Ag" "Mining" "Constr" "Manuf"

[6] "Manuf\_nd" "Transp" "Comm" "Energy" "TradeW"

[11] "TradeR" "RE" "Services" "Govt"

> X <- cbind(Ag, Mining, Constr, Manuf, Manuf\_nd, Transp, Comm,

+ Energy, TradeW, TradeR, RE, Services, Govt)

> # mean,median,25th and 75th quartiles, min, max

> # princomp performs a principal components analysis on the given

> #numeric data matrix and #returns the results as an object of

> #class princomp.

> pcal<-princomp(X, scores=TRUE, cor=TRUE)

> #Summary. A very useful multipurpose function in R issummary(X),

> #where X can be one of #any number of objects, including datasets,

> #variables, and linear models, just to name a few. #When used,

> #the command provides summary data related to the individual

> #object that was fed #into it

> summary(pcal)

Importance of components:

Comp.1 Comp.2 Comp.3

Standard deviation 1.7987525 1.4954801 1.3999420

Proportion of Variance 0.2488854 0.1720354 0.1507567

Cumulative Proportion 0.2488854 0.4209209 0.5716776

Comp.4 Comp.5 Comp.6

Standard deviation 1.1663403 1.07583525 0.93184458

Proportion of Variance 0.1046423 0.08903242 0.06679495

Cumulative Proportion 0.6763199 0.76535232 0.83214726

Comp.7 Comp.8 Comp.9

Standard deviation 0.85116719 0.78471605 0.5641253

Proportion of Variance 0.05572966 0.04736764 0.0244798

Cumulative Proportion 0.88787692 0.93524456 0.9597244

Comp.10 Comp.11 Comp.12

Standard deviation 0.4851322 0.38943836 0.36945813

Proportion of Variance 0.0181041 0.01166633 0.01049995

Cumulative Proportion 0.9778285 0.98949478 0.99999473

Comp.13

Standard deviation 8.279806e-03

Proportion of Variance 5.273476e-06

Cumulative Proportion 1.000000e+00

> # Extract or print loadings in factor analysis

> loadings(pcal)

Loadings:

Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6 Comp.7

Ag 0.135 0.385 0.373 0.411 0.245 0.433

Mining 0.470 -0.260 0.164 -0.276

Constr -0.393 0.257 -0.350 0.196 -0.370

Manuf -0.183 0.376 0.375 -0.147 0.111 -0.198 -0.151

Manuf\_nd 0.459 -0.465 -0.217 -0.102

Transp 0.418 -0.147 0.365 0.143 -0.169 -0.302

Comm -0.152 -0.316 0.343 -0.550 0.267 -0.145

Energy 0.247 0.138 -0.416 -0.202 0.689 0.199

TradeW -0.315 0.290 0.442 0.353 -0.254

TradeR -0.261 0.507 -0.227 -0.252 -0.143 0.146

RE -0.363 -0.447 0.173 0.359

Services -0.380 -0.384 -0.127 -0.183 0.125 0.101 -0.133

Govt 0.289 -0.369 -0.295 -0.306 0.428

Comp.8 Comp.9 Comp.10 Comp.11 Comp.12 Comp.13

Ag 0.277 0.152 0.217 0.286 0.206

Mining -0.148 0.116 -0.475 0.282 0.500

Constr 0.499 0.371 -0.141 -0.256

Manuf -0.500 0.387 0.138 0.406

Manuf\_nd 0.592 -0.102 0.142 0.122 0.338

Transp -0.264 0.108 0.507 -0.407 0.144

Comm -0.101 0.436 0.383

Energy -0.116 0.282 -0.286

TradeW -0.253 -0.455 -0.322 -0.210 0.112

TradeR -0.100 -0.515 0.432 -0.181 0.106

RE 0.106 0.169 -0.127 -0.496 0.451

Services -0.222 -0.204 0.458 0.460 0.320

Govt -0.121 0.171 -0.548 0.238

Comp.1 Comp.2 Comp.3 Comp.4 Comp.5 Comp.6

SS loadings 1.000 1.000 1.000 1.000 1.000 1.000

Proportion Var 0.077 0.077 0.077 0.077 0.077 0.077

Cumulative Var 0.077 0.154 0.231 0.308 0.385 0.462

Comp.7 Comp.8 Comp.9 Comp.10 Comp.11 Comp.12

SS loadings 1.000 1.000 1.000 1.000 1.000 1.000

Proportion Var 0.077 0.077 0.077 0.077 0.077 0.077

Cumulative Var 0.538 0.615 0.692 0.769 0.846 0.923

Comp.13

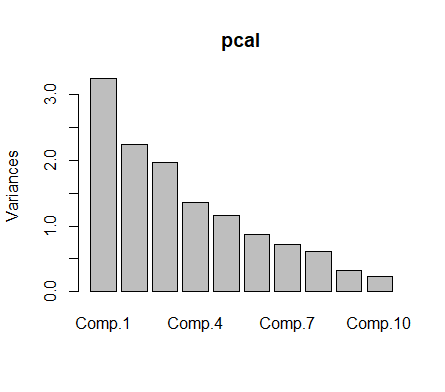
SS loadings 1.000

Proportion Var 0.077

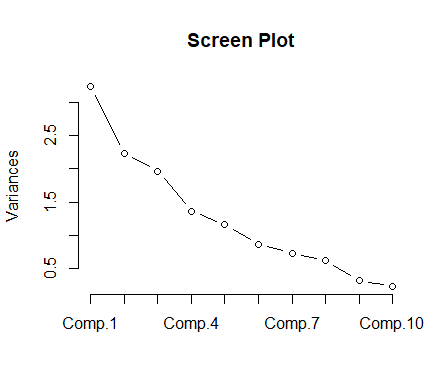
Cumulative Var 1.000

> #Visualize the Principal Components

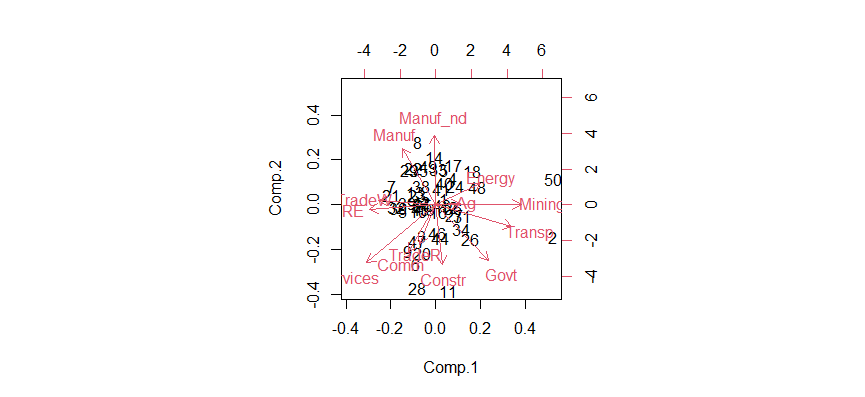
> plot(pcal)



> screeplot(pcal,type="line",main="Screen Plot")



> biplot(pcal)



> pcal$scores[1:10,]

Comp.1 Comp.2 Comp.3 Comp.4

[1,] 0.4896329 0.2840153 0.92052695 -0.08785947

[2,] 6.6852235 -1.5422261 -2.72534093 2.09383081

[3,] -0.7517456 -1.4884587 0.86440944 -1.23947821

[4,] 1.0000343 1.2527044 1.79705334 0.15952525

[5,] -1.8143951 -0.3083173 -1.07282748 0.62096975

[6,] -1.1231352 -2.7966578 -0.12662810 0.83439457

[7,] -2.4543289 0.8838887 -2.26111646 0.17696795

[8,] -0.9918322 2.9538499 -4.02639673 -0.22036149

[9,] -1.5476850 -2.2030680 0.14284640 -0.17108744

[10,] -0.9920176 -0.2998176 0.05468377 2.18146390

Comp.5 Comp.6 Comp.7 Comp.8

[1,] -1.7536053 -0.25729127 0.3122104 0.1332644

[2,] 0.9621524 -2.42578094 -1.2416981 -0.8350926

[3,] 0.7013320 -0.40238064 -0.2571765 -0.1730944

[4,] -0.6286758 0.31143230 0.2804304 0.2515513

[5,] 0.5873437 -0.19317967 0.3766092 -0.6847873

[6,] -1.5916798 0.58129969 -1.0428510 -0.1619501

[7,] 1.0153673 0.09693667 0.2158679 -0.5350483

[8,] 0.4781388 -1.24008960 1.5630880 2.3471516

[9,] -0.1473670 0.46797314 0.1884953 -0.0521831

[10,] -1.8755534 0.87740143 -0.9428300 0.1629925

Comp.9 Comp.10 Comp.11 Comp.12

[1,] 0.16656810 -0.12841084 -0.14054526 0.11795741

[2,] -0.02638485 0.03282346 0.60577408 -0.42502678

[3,] 0.13204988 0.33229286 -0.67623158 -0.53472823

[4,] -0.26955035 0.91416815 0.62811995 -0.03117104

[5,] -0.60181782 0.15025628 -0.05043902 0.57785734

[6,] 1.59879806 1.52538907 -0.02066718 0.20311745

[7,] 0.06689563 0.09648740 -0.09564494 -0.11018733

[8,] 0.38662079 0.46928579 -0.24839536 -0.84153331

[9,] -1.26973408 0.64521367 -0.42080320 -0.29586008

[10,] 0.11552867 0.08515134 -0.14312824 -0.31799226

Comp.13

[1,] -0.0022081279

[2,] 0.0038008406

[3,] -0.0060317043

[4,] 0.0028593823

[5,] -0.0072707745

[6,] -0.0025178811

[7,] 0.0001849203

[8,] -0.0004708541

[9,] -0.0079586042

[10,] -0.0020137002

Conclusion: Thus, we implementedPrincipal Component Analysis-Finding Principal Components, Variance and Standard Deviation calculations of principal components.