PCA

Abhirup Sen

20/05/2021

```
mypca <- prcomp(mtcars, scale = TRUE, center = TRUE, retx = T)
```

```
summary(mypca)
```

```
## Importance of components:
##                           PC1    PC2     PC3     PC4     PC5     PC6    PC7
## Standard deviation     2.5707 1.6280 0.79196 0.51923 0.47271 0.46000 0.3678
## Proportion of Variance 0.6008 0.2409 0.05702 0.02451 0.02031 0.01924 0.0123
## Cumulative Proportion  0.6008 0.8417 0.89873 0.92324 0.94356 0.96279 0.9751
##                           PC8    PC9    PC10   PC11
## Standard deviation     0.35057 0.2776 0.22811 0.1485
## Proportion of Variance 0.01117 0.0070 0.00473 0.0020
## Cumulative Proportion  0.98626 0.9933 0.99800 1.0000
```

```
mypca$rotation
```

```
##             PC1         PC2         PC3          PC4         PC5         PC6
## mpg  -0.3625305  0.01612440 -0.22574419 -0.022540255  0.10284468 -0.10879743
## cyl   0.3739160  0.04374371 -0.17531118 -0.002591838  0.05848381  0.16855369
## disp  0.3681852 -0.04932413 -0.06148414  0.256607885  0.39399530 -0.33616451
## hp    0.3300569  0.24878402  0.14001476 -0.067676157  0.54004744  0.07143563
## drat -0.2941514  0.27469408  0.16118879  0.854828743  0.07732727  0.24449705
## wt    0.3461033 -0.14303825  0.34181851  0.245899314 -0.07502912 -0.46493964
## qsec -0.2004563 -0.46337482  0.40316904  0.068076532 -0.16466591 -0.33048032
## vs   -0.3065113 -0.23164699  0.42881517 -0.214848616  0.59953955  0.19401702
## am   -0.2349429  0.42941765 -0.20576657 -0.030462908  0.08978128 -0.57081745
## gear -0.2069162  0.46234863  0.28977993 -0.264690521  0.04832960 -0.24356284
## carb  0.2140177  0.41357106  0.52854459 -0.126789179 -0.36131875  0.18352168
##               PC7          PC8          PC9         PC10         PC11
## mpg   0.367723810 -0.754091423  0.235701617  0.13928524 -0.124895628
## cyl   0.057277736 -0.230824925  0.054035270 -0.84641949 -0.140695441
## disp  0.214303077  0.001142134  0.198427848  0.04937979  0.660606481
## hp   -0.001495989 -0.222358441 -0.575830072  0.24782351 -0.256492062
## drat  0.021119857  0.032193501 -0.046901228 -0.10149369 -0.039530246
## wt   -0.020668302 -0.008571929  0.359498251  0.09439426 -0.567448697
## qsec  0.050010522 -0.231840021 -0.528377185 -0.27067295  0.181361780
## vs   -0.265780836  0.025935128  0.358582624 -0.15903909  0.008414634
## am   -0.587305101 -0.059746952 -0.047403982 -0.17778541  0.029823537
## gear  0.605097617  0.336150240 -0.001735039 -0.21382515 -0.053507085
## carb -0.174603192 -0.395629107  0.170640677  0.07225950  0.319594676
```

```
dim(mypca$x)
```

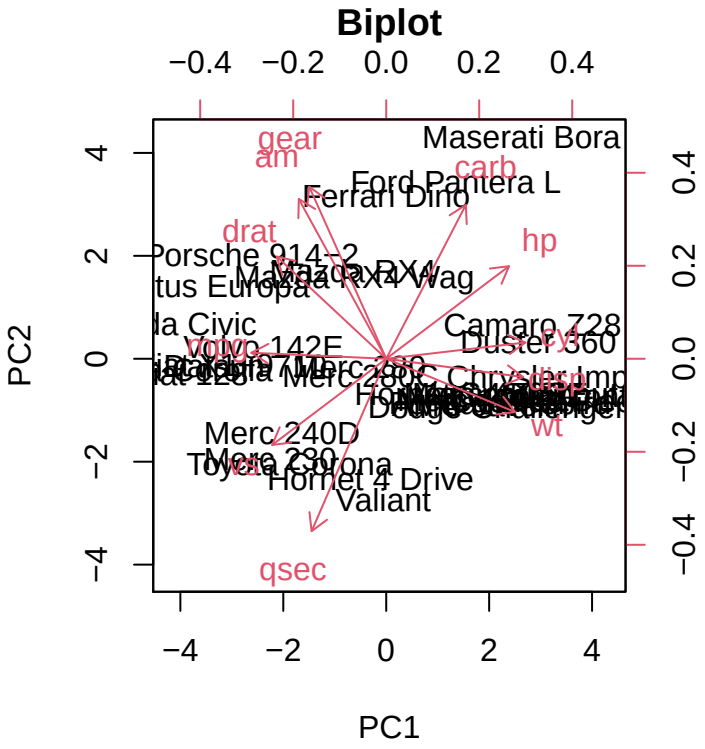
```
## [1] 32 11
```

```
mypca$x
```

```
##                           PC1        PC2        PC3          PC4
## Mazda RX4           -0.6468627420  1.7081142 -0.5917309  0.113702214
## Mazda RX4 Wag       -0.6194831460  1.5256219 -0.3763013  0.199121210
## Datsun 710          -2.7356242748 -0.1441501 -0.2374391 -0.245215450
## Hornet 4 Drive      -0.3068606268 -2.3258038 -0.1336213 -0.503800355
## Hornet Sportabout    1.9433926844 -0.7425211 -1.1165366  0.074461963
## Valiant             -0.0552534228 -2.7421229  0.1612456 -0.975167425
## Duster 360           2.9553851233  0.3296133 -0.3570461 -0.051529216
## Merc 240D           -2.0229593244 -1.4421056  0.9290295 -0.142129082
## Merc 230            -2.2513839535 -1.9522879  1.7689364  0.287210957
## Merc 280            -0.5180912217 -0.1594610  1.4692603  0.066263362
## Merc 280C           -0.5011860079 -0.3187934  1.6570701  0.094357222
## Merc 450SE           2.2124096339 -0.6727099 -0.3694707 -0.129797905
## Merc 450SL           2.0155715693 -0.6724606 -0.4768341 -0.210991001
## Merc 450SLC          2.1147047372 -0.7891129 -0.2904620 -0.175332868
## Cadillac Fleetwood   3.8383725118 -0.8149087  0.6370972  0.290505877
## Lincoln Continental  3.8918495626 -0.7218314  0.7092612  0.405336898
## Chrysler Imperial    3.5363862158 -0.4145024  0.5402468  0.665665306
## Fiat 128            -3.7955510831 -0.2920783 -0.4161681  0.055191058
## Honda Civic         -4.1870356784  0.6775721 -0.2035831  1.167526096
## Toyota Corolla      -4.1675359344 -0.2748890 -0.4589124  0.183313028
## Toyota Corona       -1.8741790870 -2.0864529  0.1543265  0.050514126
## Dodge Challenger     2.1504414942 -0.9982442 -1.1503639 -0.584982249
## AMC Javelin          1.8340369797 -0.8921886 -0.9472872  0.005694071
## Camaro Z28           2.8434957523  0.6701037 -0.1605593  0.814340105
## Pontiac Firebird     2.2105479148 -0.8600504 -1.0279577  0.146420497
## Fiat X1-9           -3.5176818134 -0.1192950 -0.4464716 -0.013427353
## Porsche 914-2       -2.6095003965  2.0141425 -0.8172519  0.568564789
## Lotus Europa        -3.3323844512  1.3568877 -0.4467167 -1.153197531
## Ford Pantera L       1.3513346957  3.4448780 -0.1343943  0.590098358
## Ferrari Dino        -0.0009743305  3.1683750  0.3957610 -0.938933017
## Maserati Bora        2.6270897605  4.3107016  1.3315940 -0.877332804
## Volvo 142E          -2.3824711412  0.2299603  0.4052798  0.223549117
##                           PC5           PC6         PC7           PC8
## Mazda RX4           -0.945523363 -0.0169873733 -0.42648652 -0.009631217
## Mazda RX4 Wag       -1.016680740 -0.2417246434 -0.41620046 -0.084520213
## Datsun 710           0.398762288 -0.3487678138 -0.60884146  0.585255765
## Hornet 4 Drive       0.549208936  0.0192969984 -0.04036075 -0.049583029
## Hornet Sportabout    0.207515698  0.1491927606  0.38350816 -0.160297757
## Valiant              0.211665375 -0.2438358546 -0.29464160  0.256612420
## Duster 360           0.343847875  0.7126920868 -0.13607792 -0.171103449
## Merc 240D           -0.316651386 -0.0009889391  0.63946214  0.163156195
## Merc 230            -0.333682355 -0.3338703384  0.62201034 -0.105779936
## Merc 280            -0.069624161  0.8165308365  0.16117090  0.099983313
## Merc 280C           -0.148803650  0.7308383757  0.09254430  0.197306566
## Merc 450SE          -0.378611141  0.1317014762 -0.01645498 -0.194092435
```

```
## Merc 450SL          -0.355611763  0.2400263805  0.05123623 -0.329669990
## Merc 450SLC         -0.432140303  0.1801997325 -0.06675316 -0.119252582
## Cadillac Fleetwood  -0.048245223 -0.8844735483 -0.16615296  0.138398783
## Lincoln Continental  0.003899176 -0.8625868981 -0.19250873  0.129305868
## Chrysler Imperial    0.208027112 -0.6536447300  0.03449804 -0.391104141
## Fiat 128             0.219981109 -0.4675796343 -0.03749941 -0.625278746
## Honda Civic          0.097674091  0.5180554279 -0.25316291 -0.395045565
## Toyota Corolla       0.222152228 -0.3171521124  0.06617540 -0.853947085
## Toyota Corona        0.039299002  0.7236992559 -0.28027808  0.207237627
## Dodge Challenger    -0.226237802  0.1062181942  0.09489585  0.316055390
## AMC Javelin         -0.252565496  0.2888101997  0.08161916  0.321900593
## Camaro Z28           0.389118986  0.9468795171 -0.21157976  0.038657331
## Pontiac Firebird     0.299261925 -0.1983310387  0.47269865 -0.234144182
## Fiat X1-9            0.206753365 -0.1449905641 -0.35850305  0.089109764
## Porsche 914-2       -0.597313744 -0.3394265065  0.82032965  0.634987241
## Lotus Europa         0.694667640  0.0165037718  0.51018011  0.004140777
## Ford Pantera L       1.101648091 -0.1746156635  0.41358868  0.609167214
## Ferrari Dino        -0.848833976 -0.0097569921  0.02967883  0.014187801
## Maserati Bora        0.455265189 -0.0156094416 -0.18813730 -0.558646792
## Volvo 142E           0.321777017 -0.3263029217 -0.77995741  0.476634473
##                             PC9         PC10         PC11
## Mazda RX4            0.14642303 -0.06670350  0.179693570
## Mazda RX4 Wag        0.07452829 -0.12692766  0.088644265
## Datsun 710          -0.13122859  0.04573787 -0.094632914
## Hornet 4 Drive       0.22021812 -0.06039981  0.147611269
## Hornet Sportabout   -0.02117623 -0.05983003  0.146406899
## Valiant             -0.03222907 -0.20165466  0.019545064
## Duster 360          -0.17844547  0.36086641  0.171863162
## Merc 240D            0.37698418  0.29086529 -0.019090358
## Merc 230            -0.86455356 -0.11597058  0.159688512
## Merc 280             0.54092449 -0.22093750 -0.124486227
## Merc 280C            0.30876072 -0.34417564 -0.034578568
## Merc 450SE          -0.05614966 -0.06531727 -0.396445135
## Merc 450SL          -0.20501055 -0.10761308 -0.197616838
## Merc 450SLC         -0.38704169 -0.21191036 -0.142498830
## Cadillac Fleetwood   0.19333387  0.06184979  0.262886205
## Lincoln Continental  0.19523562  0.12094849  0.039191100
## Chrysler Imperial    0.27447514  0.27588169 -0.224420191
## Fiat 128             0.10550311 -0.02717077 -0.208865888
## Honda Civic          0.23711675 -0.15433928  0.246835364
## Toyota Corolla      -0.11313627 -0.12606845 -0.031747839
## Toyota Corona       -0.44646972  0.51147635  0.063679725
## Dodge Challenger     0.10435633 -0.13641143  0.049594456
## AMC Javelin         -0.12237636 -0.29628634  0.045293027
## Camaro Z28          -0.05282991  0.32624525 -0.099386307
## Pontiac Firebird     0.20849043  0.01547674  0.122593248
## Fiat X1-9           -0.02228967 -0.08414018 -0.005746448
## Porsche 914-2       -0.12999660  0.34968156 -0.111596656
## Lotus Europa         0.29680350  0.23980308  0.030015592
## Ford Pantera L      -0.23280792 -0.50262890 -0.042242570
## Ferrari Dino         0.09813571  0.14491815  0.043006835
## Maserati Bora       -0.34081133  0.04706368  0.062135486
## Volvo 142E          -0.04473670  0.11767108 -0.145329008
```

```
biplot(mypca, main ="Biplot", scale=0)
```




```
mypca$sdev
```

```
##  [1] 2.5706809 1.6280258 0.7919579 0.5192277 0.4727061 0.4599958 0.3677798
##  [8] 0.3505730 0.2775728 0.2281128 0.1484736
```

#Calclate the varince Vat = Sd^2

```
mypca.var <-mypca$sdev^2
mypca.var
```

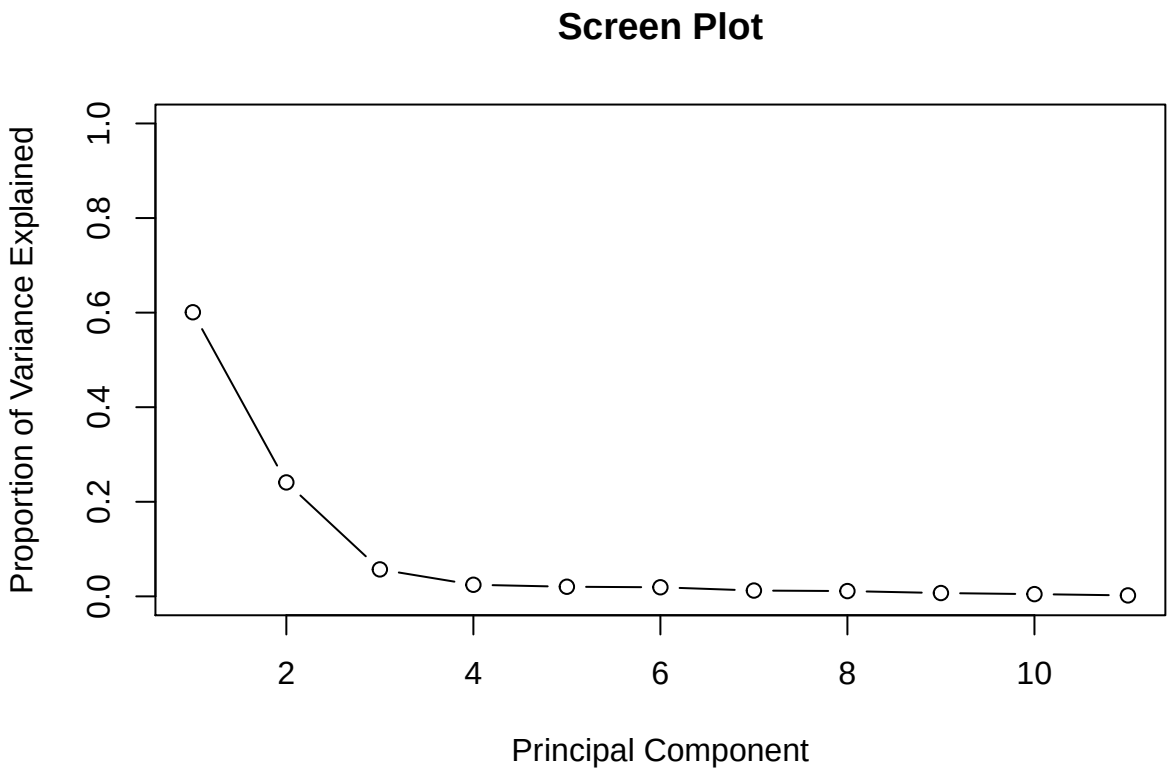
```
##  [1] 6.60840025 2.65046789 0.62719727 0.26959744 0.22345110 0.21159612
##  [7] 0.13526199 0.12290143 0.07704665 0.05203544 0.02204441
```

#proportion of Variance

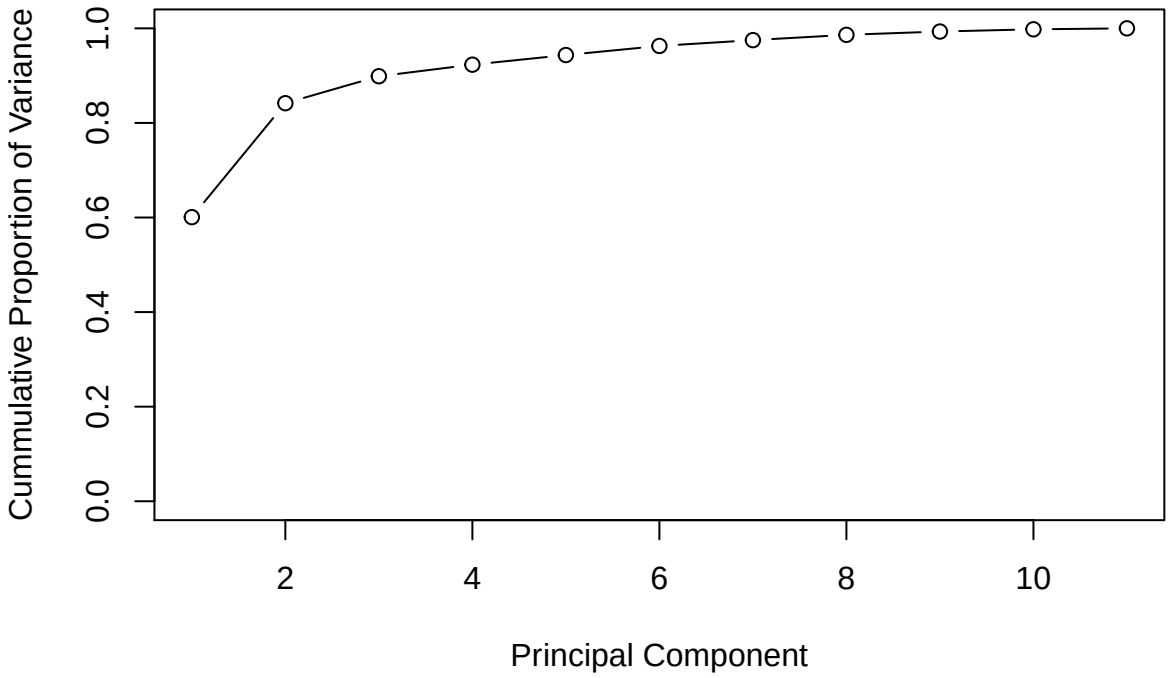
```
pve <- mypca.var / sum(mypca.var)
pve
```

```
##  [1] 0.600763659 0.240951627 0.057017934 0.024508858 0.020313737 0.019236011
##  [7] 0.012296544 0.011172858 0.007004241 0.004730495 0.002004037
```

```r
#plot the variance of each principal components
plot(pve, xlab = "Principal Component", ylab = "Proportion of Variance Explained", ylim =c(0,1),type ="
```




```r
#plot the cumulative proportion of variance
plot(cumsum(pve),xlab = "Principal Component", ylab = "Cummulative Proportion of Variance", ylim = c(0,
```




```
cumsum(pve)
```

```
##  [1] 0.6007637 0.8417153 0.8987332 0.9232421 0.9435558 0.9627918 0.9750884
##  [8] 0.9862612 0.9932655 0.9979960 1.0000000
```

```
which(cumsum(pve)>= 0.9)[1]
```

```
## [1] 4
```

#lets now train the model #add the training model # we ill predict mileage

```
train.data <- data.frame(mpg = mtcars$mpg, mypca$x[,1:4])
```

#lets run a decision tree

```
install.packages("rpart", repos="http://cran.us.r-project.org")
```

```
## Installing package into 'C:/Users/LENOVO/OneDrive/Documents/R/win-library/4.0'
## (as 'lib' is unspecified)
```

```
## package 'rpart' successfully unpacked and MD5 sums checked
```

```
## Warning: cannot remove prior installation of package 'rpart'
```

```
## Warning in file.copy(savedcopy, lib, recursive = TRUE):
## problem copying C:\Users\LENOVO\OneDrive\Documents\R\win-
## library\4.0\00LOCK\rpart\libs\x64\rpart.dll to C:
## \Users\LENOVO\OneDrive\Documents\R\win-library\4.0\rpart\libs\x64\rpart.dll:
## Permission denied

## Warning: restored 'rpart'

##
## The downloaded binary packages are in
##  C:\Users\LENOVO\AppData\Local\Temp\Rtmp8absF8\downloaded_packages
```

```
library(rpart)
```

```
rpart.model<- rpart(mpg ~ ., data=train.data  , method ="anova")
```

```
library(rpart.plot)
rpart.plot(rpart.model)
```