tmp.R

Velociraptor

Sat Jun 10 20:35:41 2017

```
require(graphics)

## The variances of the variables in the
## USArrests data vary by orders of magnitude, so scaling is appropriate
(pc.cr <- princomp(USArrests))  # inappropriate
```

```
## Call:
## princomp(x = USArrests)
##
## Standard deviations:
##    Comp.1    Comp.2    Comp.3    Comp.4
## 82.890847 14.069560  6.424204  2.457837
##
##  4  variables and  50 observations.
```

```
princomp(USArrests, cor = TRUE) # =^= prcomp(USArrests, scale=TRUE)
```

```
## Call:
## princomp(x = USArrests, cor = TRUE)
##
## Standard deviations:
##    Comp.1    Comp.2    Comp.3    Comp.4
## 1.5748783 0.9948694 0.5971291 0.4164494
##
##  4  variables and  50 observations.
```

```
## Similar, but different:
## The standard deviations differ by a factor of sqrt(49/50)

summary(pc.cr <- princomp(USArrests, cor = TRUE))
```

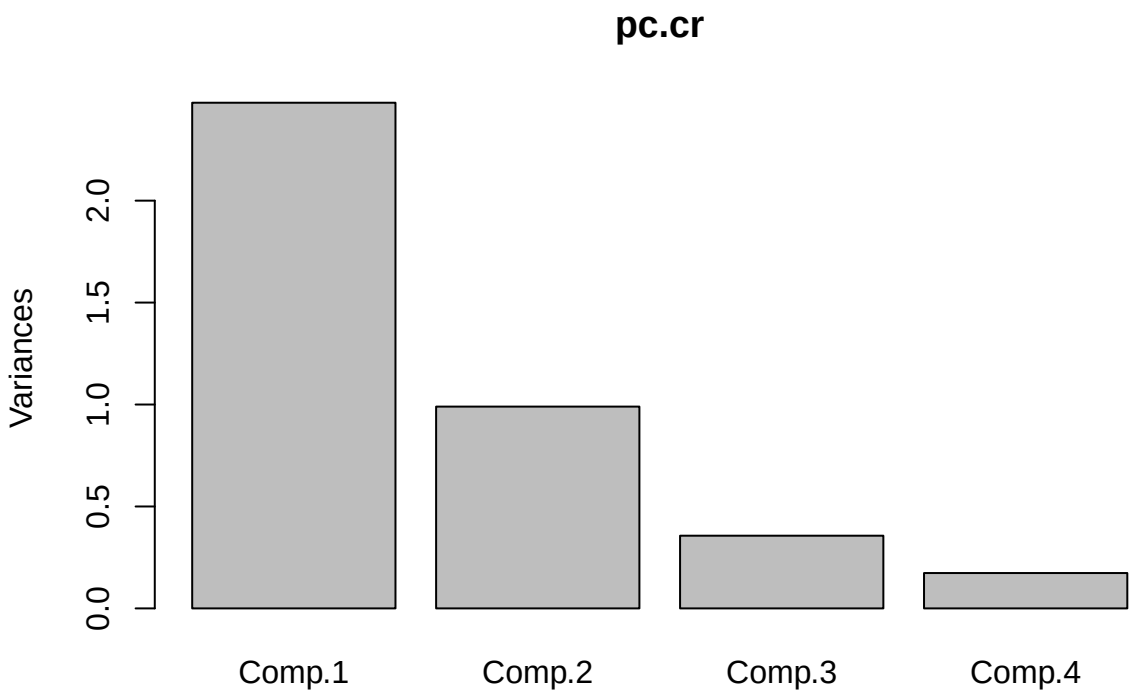
```
## Importance of components:
##                            Comp.1    Comp.2    Comp.3     Comp.4
## Standard deviation     1.5748783 0.9948694 0.5971291 0.41644938
## Proportion of Variance 0.6200604 0.2474413 0.0891408 0.04335752
## Cumulative Proportion  0.6200604 0.8675017 0.9566425 1.00000000
```

```
loadings(pc.cr)  # note that blank entries are small but not zero
```

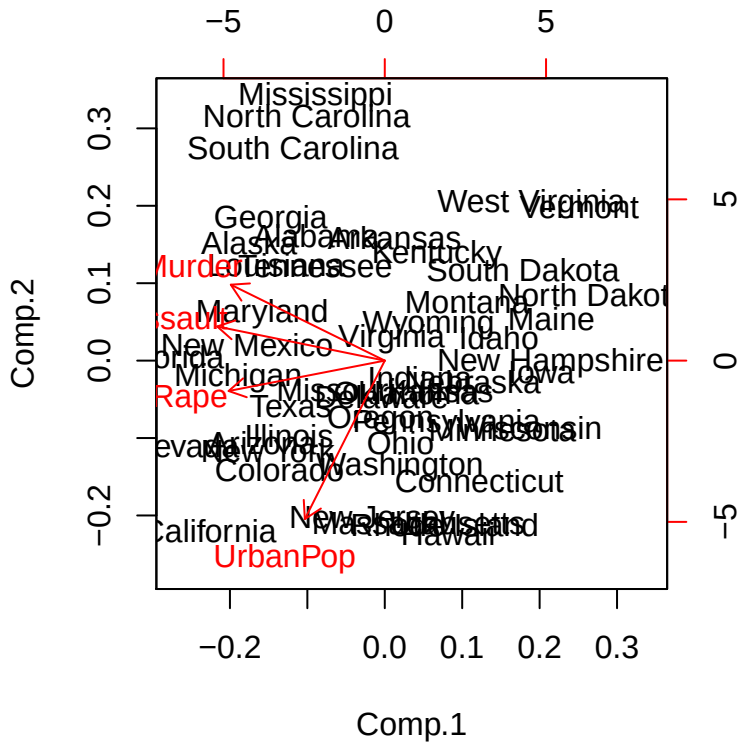
```
##
## Loadings:
##          Comp.1 Comp.2 Comp.3 Comp.4
## Murder   -0.536  0.418 -0.341  0.649
## Assault  -0.583  0.188 -0.268 -0.743
## UrbanPop -0.278 -0.873 -0.378  0.134
## Rape     -0.543 -0.167  0.818
##
##                Comp.1 Comp.2 Comp.3 Comp.4
## SS loadings      1.00   1.00   1.00   1.00
```

```
## Proportion Var   0.25   0.25   0.25   0.25
## Cumulative Var   0.25   0.50   0.75   1.00
```

```
## The signs of the columns are arbitrary
plot(pc.cr) # shows a screeplot.
```




```
biplot(pc.cr)
```




```
## Formula interface
princomp(~ ., data = USArrests, cor = TRUE)
```

```
## Call:
## princomp(formula = ~., data = USArrests, cor = TRUE)
##
## Standard deviations:
##    Comp.1    Comp.2    Comp.3    Comp.4
## 1.5748783 0.9948694 0.5971291 0.4164494
##
##  4  variables and  50 observations.
```

```
## NA-handling
USArrests[1, 2] <- NA
pc.cr <- princomp(~ Murder + Assault + UrbanPop,
                  data = USArrests, na.action = na.exclude, cor = TRUE)
pc.cr$scores[1:5, ]
```

```
##                 Comp.1     Comp.2     Comp.3
## Alabama             NA         NA         NA
## Alaska     -0.80197919 -1.4204705 -0.6423222
## Arizona    -1.38652787  0.7860977 -0.8455448
## Arkansas   -0.04279066 -1.1307711 -0.1791254
## California -1.58764207  1.4584727 -0.4217155
```

```
## (Simple) Robust PCA:
## Classical:
(pc.cl  <- princomp(stackloss))
```

```
## Call:
## princomp(x = stackloss)
##
## Standard deviations:
##    Comp.1    Comp.2    Comp.3    Comp.4
## 13.596589  4.676077  2.617533  1.366320
##
##  4  variables and  21 observations.
```

```
## Robust:
(pc.rob <- princomp(stackloss, covmat = MASS::cov.rob(stackloss)))
```

```
## Call:
## princomp(x = stackloss, covmat = MASS::cov.rob(stackloss))
##
## Standard deviations:
##    Comp.1    Comp.2    Comp.3    Comp.4
## 7.8322873 4.0077676 1.9114016 0.7624211
##
##  4  variables and  21 observations.
```