9.R

AVShrez

2024-05-31

```
# Creating a flower scale using PCA

library(dplyr)
```

```
##
## Attaching package: 'dplyr'
## The following objects are masked from 'package:stats':
##
##     filter, lag
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
flower.scale <- iris[,-5] %>% scale

flowerpca <- prcomp(flower.scale)

flowerpca
```

```
## Standard deviations (1, .., p=4):
## [1] 1.7083611 0.9560494 0.3830886 0.1439265
##
## Rotation (n x k) = (4 x 4):
##                     PC1         PC2        PC3        PC4
## Sepal.Length  0.5210659 -0.37741762  0.7195664  0.2612863
## Sepal.Width  -0.2693474 -0.92329566 -0.2443818 -0.1235096
## Petal.Length  0.5804131 -0.02449161 -0.1421264 -0.8014492
## Petal.Width   0.5648565 -0.06694199 -0.6342727  0.5235971
```

```
summary(flowerpca)
```

```
## Importance of components:
##                           PC1    PC2     PC3     PC4
## Standard deviation     1.7084 0.9560 0.38309 0.14393
## Proportion of Variance 0.7296 0.2285 0.03669 0.00518
## Cumulative Proportion  0.7296 0.9581 0.99482 1.00000
```

```
# Computing EigenValues

eigenvalues.PC1 <- flowerpca$sdev[1]^2
eigenvalues.PC2 <- flowerpca$sdev[2]^2
eigenvalues.PC3 <- flowerpca$sdev[3]^2
eigenvalues.PC4 <- flowerpca$sdev[4]^2
```

```r
eigenV <- data.frame(eigenvalues.PC1, eigenvalues.PC2, eigenvalues.PC3, eigenvalues.PC4)
eigenV
```

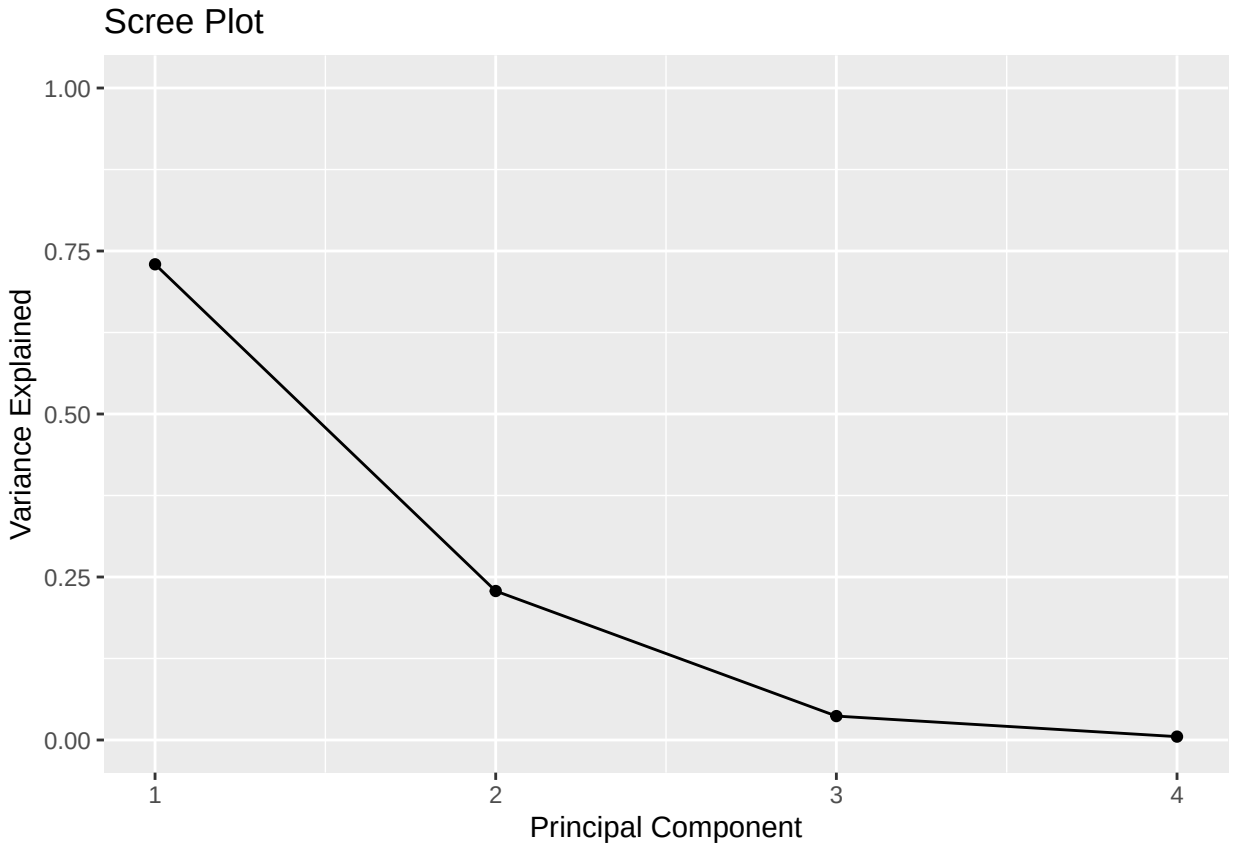
```
##   eigenvalues.PC1 eigenvalues.PC2 eigenvalues.PC3 eigenvalues.PC4
## 1        2.918498       0.9140305       0.1467569      0.02071484
```

```r
# Kaiser's Rule states that Principal Component with Eigenvalue >= 1 must be retained for latent variab
# So here, PC1 is the only component required to create the latent variable.

# Scree Plot
# calculating total variance explained by each principal component
varExplained = flowerpca$sdev^2 / sum(flowerpca$sdev^2)

varExplained
```

```
## [1] 0.729624454 0.228507618 0.036689219 0.005178709
```

```r
#create scree plot
library(ggplot2)
qplot(c(1:4), varExplained) +
  geom_line() +
  xlab("Principal Component") +
  ylab("Variance Explained") +
  ggtitle("Scree Plot") +
  ylim(0, 1)
```

```
## Warning: `qplot()` was deprecated in ggplot2 3.4.0.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warning was
## generated.
```




```
# So here, PC1 is the only component required to be retained as it explains the 72% variance.

# with VARIMAX rotation
# flowerpca2 <- psych::principal(flowerpca, nfactors = 3, rotate = "varimax")
# summary(flowerpca2)
```