Assignment 2 DSA6100

Shrankhala Jain

1/22/2018

Applying PCA on the Iris flower dataset for analyzing and visualization. The data contain four continuous variables which corresponds to physical measures of flowers and a categorical variable describing the flowers’ species. Used prcomp package and ggbiplot for the visualization.

## R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see <http://rmarkdown.rstudio.com>.

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

log.ir<- log(iris[, 1:4])  
ir.species <- iris[, 5]  
ir.pca <- prcomp(log.ir,  
 center = TRUE,  
 scale. = TRUE)   
print(ir.pca)

## Standard deviations (1, .., p=4):  
## [1] 1.7124583 0.9523797 0.3647029 0.1656840  
##   
## Rotation (n x k) = (4 x 4):  
## PC1 PC2 PC3 PC4  
## Sepal.Length 0.5038236 -0.45499872 0.7088547 0.19147575  
## Sepal.Width -0.3023682 -0.88914419 -0.3311628 -0.09125405  
## Petal.Length 0.5767881 -0.03378802 -0.2192793 -0.78618732  
## Petal.Width 0.5674952 -0.03545628 -0.5829003 0.58044745

## Including Plots

## Warning: package 'devtools' was built under R version 3.4.2

## Warning: Username parameter is deprecated. Please use vqv/ggbiplot

## Warning in strptime(x, fmt, tz = "GMT"): unknown timezone 'zone/tz/2017c.  
## 1.0/zoneinfo/America/Detroit'

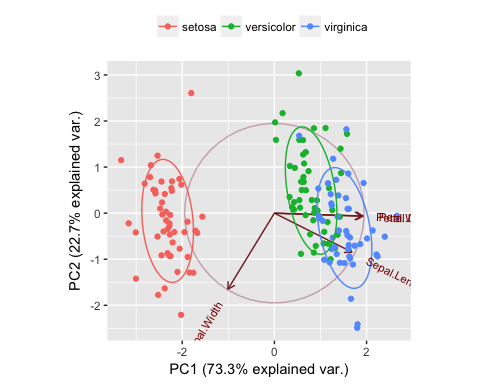
## Skipping install of 'ggbiplot' from a github remote, the SHA1 (7325e880) has not changed since last install.  
## Use `force = TRUE` to force installation

## Loading required package: ggplot2

## Loading required package: plyr

## Loading required package: scales

## Loading required package: grid

 The plot clearly indicates that one class can be separated from the other two those are merging together.

Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.