Responsi - PPJ 21/04/2020 - 3SK3 - APG.R

Ridson Alfarizal

2022-04-21

**library**(tidyverse)

print\_eq <- function(eig\_vec, comp = 1:ncol(eig\_vec)){  
 for (i in comp) {  
 cat(stringr::str\_glue('y{i} = {paste0(round(eig\_vec[,i], 3), "\*X", 1:length(eig\_vec[,i]), collapse = " + ")}'), '\n')  
 }  
}  
scree\_plot <- function(eig\_values){  
 ggplot(NULL, aes(y = eig\_values, x = 1:length(eig\_values))) +  
 geom\_line() +  
 geom\_point(color = 'red') +  
 geom\_text(aes(label = round(eig\_values, 3)), nudge\_x = 0.2, nudge\_y = 0.2) +  
 labs(  
 x = 'Component Number',  
 y = 'Eigen Value',  
 title = 'Scree Plot'  
 ) +  
 theme\_bw()  
}

R <- matrix(c(1, 0.9, 0.74, 0.64,  
 0.9, 1, 0.33, 0.74,  
 0.74, 0.33, 1, 0.66,  
 0.64, 0.74, 0.66, 1),  
 nrow = 4, byrow = T)  
R

## [,1] [,2] [,3] [,4]  
## [1,] 1.00 0.90 0.74 0.64  
## [2,] 0.90 1.00 0.33 0.74  
## [3,] 0.74 0.33 1.00 0.66  
## [4,] 0.64 0.74 0.66 1.00

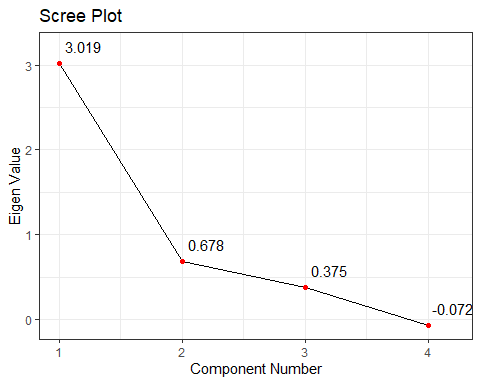
# A -------------------------------------------------------------------  
eig <- eigen(R)  
eig

## eigen() decomposition  
## $values  
## [1] 3.01878551 0.67793201 0.37519449 -0.07191201  
##   
## $vectors  
## [,1] [,2] [,3] [,4]  
## [1,] -0.5459384 0.08698788 0.56959722 -0.6082297  
## [2,] -0.5004768 0.62861750 0.05704317 0.5925446  
## [3,] -0.4461757 -0.77203698 0.14729690 0.4280068  
## [4,] -0.5023958 -0.03510062 -0.80660319 -0.3094474

# proporsi  
eig$values / sum(eig$values)

## [1] 0.75469638 0.16948300 0.09379862 -0.01797800

# scree plot  
scree\_plot(eig$values)



# maka komponen utama yang sebaiknya digunakan adalah 1  
# karena sudah dapat menjelaskan 75,47% variasi. Selain itu berdasarkan  
# scree plot, nilai eigen yang lebih besar dari satu hanya dicirikan   
# oleh 1 komponen.  
  
  
  
# B -------------------------------------------------------------------  
print\_eq(eig$vector, 1)

## y1 = -0.546\*X1 + -0.5\*X2 + -0.446\*X3 + -0.502\*X4

# Penjelasan :  
# berdasarkan hasil poin a didapat proporsinya sebesar 75.47 %,   
# yang artinya minimal dapat menangkap 75.47 % keragaman dari data  
# awal  
  
  
  
# C -------------------------------------------------------------------  
sweep(  
 eig$vectors, 2, eig$values,  
 FUN = function(a, b) a\*sqrt(b)  
) %>%   
 `dimnames<-`(  
 list(paste0('y', 1:nrow(.)), paste0('x', 1:ncol(.)))  
 )

## Warning in sqrt(b): NaNs produced

## x1 x2 x3 x4  
## y1 -0.9485489 0.07162288 0.34889608 NaN  
## y2 -0.8695611 0.51758245 0.03494072 NaN  
## y3 -0.7752147 -0.63566921 0.09022395 NaN  
## y4 -0.8728952 -0.02890067 -0.49406963 NaN

# Penjelasan :  
# berdasarkan hasil diatas variabel x1 memberikan pengaruh yang besar terhadap y1 kemudian disusul oleh variabel x3. artinya variabel y1 merupakan penciri dari faktor demografi dan ekonomi.