project5

Jason Grahn

2/16/2019

library(CCA)  
library(tidyverse)  
library(GGally)  
library(yacca)

# Use R to solve Chapter 11 Page 402: #11.9

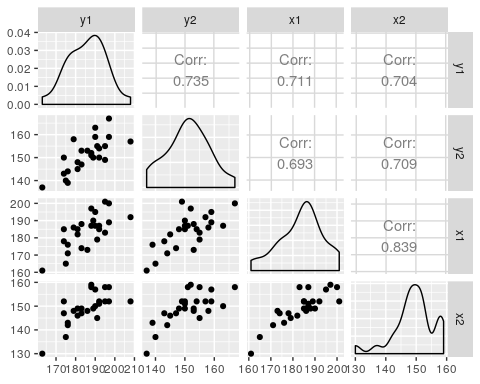
Make sure you include the commands and outputs, as well as the interpretations of the outputs.

11.9 Use the sons data of Table 3.8.

sons <- read.table(here::here("assignment06/T3\_8\_SONS.DAT")) %>%   
 rename(y1 = V1, # head length  
 y2 = V2, # head breadth  
 x1 = V3, # head length  
 x2 = V4) # head breadth  
  
knitr::kable(head(sons, 5))

|  |  |  |  |
| --- | --- | --- | --- |
| y1 | y2 | x1 | x2 |
| 191 | 155 | 179 | 145 |
| 195 | 149 | 201 | 152 |
| 181 | 148 | 185 | 149 |
| 183 | 153 | 188 | 149 |
| 176 | 144 | 171 | 142 |

ggpairs(sons)



#subgroup the data into first and second sons  
first.son <- sons %>% select(y1, y2)  
second.son <- sons %>% select(x1, x2)

cca\_output <- cca(first.son, second.son)  
cc\_output <- cc(first.son, second.son)

## (a) Find the canonical correlations between (,) and (, )·

cca\_output$corr

## CV 1 CV 2   
## 0.7885079 0.0537397

is 0.7885079 and is 0.0537397. This shows us that the canonical correlation between [one thing] and [the other thing] is positive and strong, while the canonical correlation between [something else] and [one more thing], while positive, is quite weak.

# (b) Find the standardized coefficients for the canonical variates.

cca\_output[3:4]

## $xcoef  
## CV 1 CV 2  
## y1 -0.05656620 -0.1399711  
## y2 -0.07073683 0.1869496  
##   
## $ycoef  
## CV 1 CV 2  
## x1 -0.0502426 -0.1761479  
## x2 -0.0802224 0.2620836

# (c) Test the significance of each canonical correlation.

F.test.cca(cca\_output)

##   
## F Test for Canonical Correlations (Rao's F Approximation)  
##   
## Corr F Num df Den df Pr(>F)   
## CV 1 0.788508 6.597193 4.000000 42 0.0003256 \*\*\*  
## CV 2 0.053740 0.063719 1.000000 22 0.8030550   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

The Canonical correlation value of (0.7885079) between the head length and head breadth for the *first son* tests to significance with a p-value of 0.0003256.

The output of the F-test also shows us that the canonical correlation between the head length and head breadth for the *second son* is **not** significant. This tested out to a p-value of 0.8030550 given it’s canonical correlation of 0.053740.