Canonical Correlation Analysis in Body Fat Measure

S18809

May 31, 2024

1 Introduction

Canonical correlation analysis is used to identify and measure the associations among two sets of variables. Canonical correlation is appropriate in the same situations where multiple regression would be, but where are there are multiple intercorrelated outcome variables. Canonical correlation analysis determines a set of canonical variates, orthogonal linear combinations of the variables within each set that best explain the variability both within and between sets.

Canonical correlations are invariant to changes of scale on either the response variables (y's) and the explanatory variables (x's). In other words, changing the scale of measurement of the 2 sets of variables of interest in the analysis, for instance, from inches to centimeters does not interfere with the canonical correlations that follow.

The bodyfat prediction data set is what we're utilizing to clarify this requirement.There are 252 rows and 15 columns in this data collection.Objective of the dataset is estimate of percentage of body fat determines by under Water.The data set variable listed Here:

Density - Density Determine From under water weightining
Percent body fat from Siri's (1956) equation
Age (years)
Weight (lbs)
Height (inches)
Neck circumference (cm)
Chest circumference (cm)
Abdomen 2 circumference (cm)
Hip circumference (cm)
Thigh circumference (cm)
Knee circumference (cm)
Ankle circumference (cm)
Biceps (extended) circumference (cm)
Forearm circumference (cm)
Wrist circumference (cm)

Table 1: A column description for bodyfat prediction data set

The data set was cleaned and variable separated two sets.Our goal is measured relation ship between the body _composition_metrics and body_measurements.

body_composition_matrics	Density,BodyFat,Age,Weight,Height
body_measurements	Neck,Chest,Abdomen,Hip,Thigh,Knee,Ankle,Biceps,Forearm,Wrist

Table 2: Two set of the Data

2 Methodology

2.1 Loading Appropriate Library

```
library(ggplot2)
library(GGally)
library(CCA)#facilitates canonical correlation analysis
library(CCP)#facilitates checking the significance of the canonical variates
library(dplyr)
library(tidyverse)
library(corrplot)
```

2.2 Load the Data Frame

```
bodyFat <- read_csv(file = "../Data/bodyfat.csv")
```

2.3 Define the two set

```
body_composition_metrics <- bodyFat[, 1:5] #U_ X variables
body_measurements <- bodyFat[, 6:15] #V_ Y variables
```

Select U and V data set such as body_consumption_matrics and body_measurements based on the number of variables that exist in each set so that $p(5) <= q(10)$

2.4 Find the correlation between Each variables and each data set

```
# Compute correlation matrix
correlation_matrix <- cor(bodyFat)

# Draw correlation matrix with customization
corrplot(correlation_matrix, method = "number", col = colorRampPalette(c("blue", "white",
type = "upper", order = "hclust", tl.col = "black", tl.srt = 45,
addrect = 3, rect.col = "grey", number.cex = 0.7, tl.cex = 0.8,
title = "Correlation Matrix")

ggpairs(body_composition_metrics)

ggpairs(body_measurements)
```

2.5 Checking the Correlation between Associate set

```
matcor(body_composition_metrics,body_measurements)
```

2.6 Obtain the Canonical Correlation Between Two set

```
CC_bodyFat <- cc(body_composition_metrics,body_measurements)
CC_bodyFat$cor

squared_canonical_correlations <- (CC_bodyFat$cor)^2
print("Squared Canonical Correlations:")
print(squared_canonical_correlations)

#raw canonical coefficients
CC_bodyFat[3:4]

cc_bodyFat_2 <- comput(body_composition_metrics,body_measurements,CC_bodyFat)

#displays the canonical loading
cc_bodyFat_2[3:6]
```

2.7 Test the Hypothesis testing

$H_0 : \rho_i^* = 0$
$H_1 : at\,least\,one\,\rho_i^* \neq 0 \quad \forall$ i = 1,2,3,4,5

```
rho <- CC_bodyFat$cor

n <- dim(body_composition_metrics)[1]
p <- length(body_composition_metrics)
q <- length(body_measurements)

p.asym(rho, n, p, q, tstat = "Wilks")
p.asym(rho, n, p, q, tstat = "Hotelling")
p.asym(rho, n, p, q, tstat = "Pillai")
p.asym(rho, n, p, q, tstat = "Roy")
```

2.8 Standardize the Data set

```
# Standardize first canonical variables
std_1<-diag(sqrt(diag(cov(body_composition_metrics))))
ss_1<- std_1%*%CC_bodyFat$xcoef
`rownames<-`(ss_1,c("Density","BodyFat","Age","Weight","Height"))


# Standardize second canonical variables
std_2<-diag(sqrt(diag(cov(body_measurements))))
ss_2<- std_2%*%CC_bodyFat$ycoef
`rownames<-`(ss_2,c("Neck","Chest","Abdomen","Hip","Thigh","Knee","Ankle","Biceps"
,"Forearm","Wrist"))
```

3 Result and Discussion

3.1 Correlation between Each Variables and set

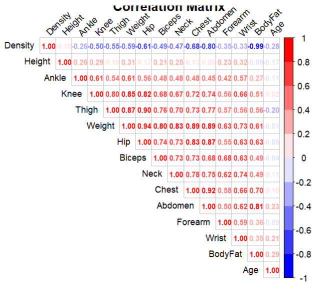


Figure 1: Figure1

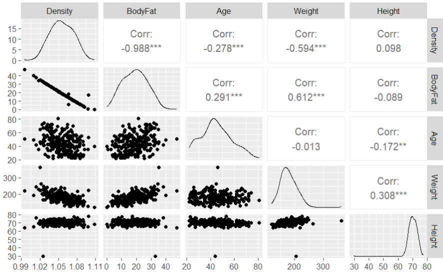


Figure 2: Figure2

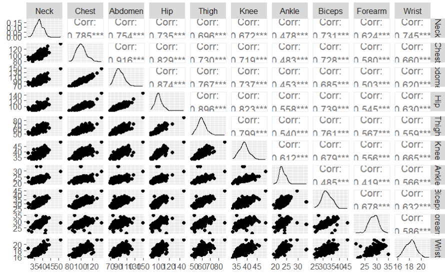


Figure 3: Figure3

Figure 1 illustrates the correlation between each variable in the data set. Here, most of the variables are highly associated with each other except Age and Height variables. The Density variable is negatively associated with all variables except Height. Also, most of the variables are negatively correlated with Age variables.

In Figure 2, we clearly represent the association between each variable in the body consumption matrix. In this case, BodyFat has a strong negative relationship with Density (-0.988) and a moderate relationship with Weight (-0.594). In this case, the BodyFat variable is a positive-storing relation to the Weight variable.

Figure 3 shows the association between each variable in the body measurement matrix. Here, we can clearly see that all variables are positive and highly correlated with each other.

3.2 Check the association Between the set

```
$Xcor
Density      BodyFat          Age       Weight       Height
Density  1.00000000 -0.98778240 -0.27763721 -0.59406188  0.09788114
BodyFat -0.98778240  1.00000000  0.29145844  0.61241400 -0.08949538
Age     -0.27763721  0.29145844  1.00000000 -0.01274609 -0.17164514
Weight  -0.59406188  0.61241400 -0.01274609  1.00000000  0.30827854
Height   0.09788114 -0.08949538 -0.17164514  0.30827854  1.00000000

$Ycor
Neck       Chest    Abdomen        Hip     Thigh       Knee      Ankle
Neck     1.0000000 0.7848350 0.7540774 0.7349579 0.6956973 0.6724050 0.4778924
Chest    0.7848350 1.0000000 0.9158277 0.8294199 0.7298586 0.7194964 0.4829879
Abdomen  0.7540774 0.9158277 1.0000000 0.8740662 0.7666239 0.7371789 0.4532227
Hip      0.7349579 0.8294199 0.8740662 1.0000000 0.8964098 0.8234726 0.5583868
Thigh    0.6956973 0.7298586 0.7666239 0.8964098 1.0000000 0.7991703 0.5397971
Knee     0.6724050 0.7194964 0.7371789 0.8234726 0.7991703 1.0000000 0.6116082
Ankle    0.4778924 0.4829879 0.4532227 0.5583868 0.5397971 0.6116082 1.0000000
Biceps   0.7311459 0.7279075 0.6849827 0.7392725 0.7614774 0.6787088 0.4848545
Forearm  0.6236603 0.5801727 0.5033161 0.5450141 0.5668422 0.5558982 0.4190500
Wrist    0.7448264 0.6601623 0.6198324 0.6300895 0.5586848 0.6645073 0.5661946
Biceps     Forearm      Wrist
Neck     0.7311459 0.6236603 0.7448264
Chest    0.7279075 0.5801727 0.6601623
Abdomen  0.6849827 0.5033161 0.6198324
Hip      0.7392725 0.5450141 0.6300895
Thigh    0.7614774 0.5668422 0.5586848
Knee     0.6787088 0.5558982 0.6645073
Ankle    0.4848545 0.4190500 0.5661946
Biceps   1.0000000 0.6782551 0.6321264
Forearm  0.6782551 1.0000000 0.5855883
Wrist    0.6321264 0.5855883 1.0000000

$XYcor
Density      BodyFat          Age       Weight       Height
Density  1.00000000 -0.98778240 -0.27763721 -0.59406188  0.09788114
BodyFat -0.98778240  1.00000000  0.29145844  0.61241400 -0.08949538
Age     -0.27763721  0.29145844  1.00000000 -0.01274609 -0.17164514
Weight  -0.59406188  0.61241400 -0.01274609  1.00000000  0.30827854
Height   0.09788114 -0.08949538 -0.17164514  0.30827854  1.00000000
```

```
Neck     -0.47296636  0.49059185  0.11350519  0.83071622  0.25370988
Chest    -0.68259865  0.70262034  0.17644968  0.89419052  0.13489181
Abdomen  -0.79895463  0.81343228  0.23040942  0.88799494  0.08781291
Hip      -0.60933143  0.62520092 -0.05033212  0.94088412  0.17039426
Thigh    -0.55309098  0.55960753 -0.20009576  0.86869354  0.14843561
Knee     -0.49504035  0.50866524  0.01751569  0.85316739  0.28605321
Ankle    -0.26489003  0.26596977 -0.10505810  0.61368542  0.26474369
Biceps   -0.48710872  0.49327113 -0.04116212  0.80041593  0.20781557
Forearm  -0.35164842  0.36138690 -0.08505555  0.63030143  0.22864922
Wrist    -0.32571598  0.34657486  0.21353062  0.72977489  0.32206533
Neck        Chest      Abdomen         Hip       Thigh        Knee
Density -0.4729664 -0.6825987 -0.79895463 -0.60933143 -0.5530910 -0.49504035
BodyFat  0.4905919  0.7026203  0.81343228  0.62520092  0.5596075  0.50866524
Age      0.1135052  0.1764497  0.23040942 -0.05033212 -0.2000958  0.01751569
Weight   0.8307162  0.8941905  0.88799494  0.94088412  0.8686935  0.85316739
Height   0.2537099  0.1348918  0.08781291  0.17039426  0.1484356  0.28605321
Neck     1.0000000  0.7848350  0.75407737  0.73495788  0.6956973  0.67240498
Chest    0.7848350  1.0000000  0.91582767  0.82941992  0.7298586  0.71949640
Abdomen  0.7540774  0.9158277  1.00000000  0.87406618  0.7666239  0.73717888
Hip      0.7349579  0.8294199  0.87406618  1.00000000  0.8964098  0.82347262
Thigh    0.6956973  0.7298586  0.76662393  0.89640979  1.0000000  0.79917030
Knee     0.6724050  0.7194964  0.73717888  0.82347262  0.7991703  1.00000000
Ankle    0.4778924  0.4829879  0.45322269  0.55838682  0.5397971  0.61160820
Biceps   0.7311459  0.7279075  0.68498272  0.73927252  0.7614774  0.67870883
Forearm  0.6236603  0.5801727  0.50331609  0.54501412  0.5668422  0.55589819
Wrist    0.7448264  0.6601623  0.61983243  0.63008954  0.5586848  0.66450729
Ankle       Biceps      Forearm       Wrist
Density -0.2648900 -0.48710872 -0.35164842 -0.3257160
BodyFat  0.2659698  0.49327113  0.36138690  0.3465749
Age     -0.1050581 -0.04116212 -0.08505555  0.2135306
Weight   0.6136854  0.80041593  0.63030143  0.7297749
Height   0.2647437  0.20781557  0.22864922  0.3220653
Neck     0.4778924  0.73114592  0.62366027  0.7448264
Chest    0.4829879  0.72790748  0.58017273  0.6601623
Abdomen  0.4532227  0.68498272  0.50331609  0.6198324
Hip      0.5583868  0.73927252  0.54501412  0.6300895
Thigh    0.5397971  0.76147745  0.56684218  0.5586848
Knee     0.6116082  0.67870883  0.55589819  0.6645073
Ankle    1.0000000  0.48485454  0.41904999  0.5661946
Biceps   0.4848545  1.00000000  0.67825513  0.6321264
Forearm  0.4190500  0.67825513  1.00000000  0.5855883
Wrist    0.5661946  0.63212642  0.58558825  1.0000000
```

This matrix illustrates the correlation between each set and the correlation matrix in XY, where X represents body_consumption_matrics and Y represents body_measurements.

3.3 Coefficient of the Canonical variables

```
$xcoef
[,1]         [,2]        [,3]         [,4]         [,5]
Density   4.275036756 12.55488437 40.84051546 193.77309883 276.23252853
BodyFat -0.003944937 -0.05957253 -0.02363106   0.37908414   0.69663308
Age     -0.003658936 -0.05199219  0.06289553   0.01146281  -0.02507464
Weight  -0.032115108  0.02070196  0.01901430   0.01615397  -0.01988686
Height   0.021905393 -0.03837303  0.09419517  -0.24872386   0.15159801

$ycoef
[,1]         [,2]         [,3]         [,4]        [,5]
Neck    -0.033731736  0.10110192  0.204602877 -0.02566358 -0.06773401
Chest   -0.022818444  0.04531884  0.027687793  0.17633820  0.04305294
Abdomen -0.024276798 -0.20118641 -0.087365019 -0.12341571 -0.03329748
Hip     -0.042365972  0.16047755  0.048835486  0.19658606  0.07769563
Thigh   -0.007017635  0.11199988 -0.203187010  0.00639431 -0.09374728
Knee    -0.038059042 -0.05946626  0.244674075 -0.33594851  0.32064061
Ankle   -0.026822001  0.09565911 -0.067146022 -0.35190756 -0.25876402
Biceps  -0.021032800 -0.01593798  0.007373403 -0.15905422 -0.34343366
Forearm -0.006994097  0.03814112 -0.152492420 -0.20623410  0.53769542
Wrist   -0.025139906 -0.15390251  0.884782812  0.35466292 -0.30242144
```

Here we can find the coefficient of each variable associated with the canonical correlation.

Interpret the x coefficient

$$U_1 = 4.2750X_{Density} + (-0.0039)X_{BodyFat} + (-0.0036)X_{Age} + (-0.032)X_{Weight} + 0.0219X_{Height} \quad (1)$$

Consider the two variables Density and BodyFat, **Density** has a positive coefficient, i.e., when we increase Density by one unit, then U_1 increases by 4.2750 units. **BodyFat** has a negative Coefficient i.e When we increased BodyFat by one unit,then U_1 decreased by 0.0039 unit. Similarly,we can interpret coefficients give the contribution of the individual variables to the First canonical Form.This way we can write other canonical variables such as U_2, U_3, U_4 and U_5

Interpret the Y coefficient

$$V_1 = -0.0337Y_{Neck}+(-0.0228)Y_{Chest}+(-0.0242)Y_{Abdoment}+(-0.042)Y_{Hip}+(-0.0070)Y_{Thigh}+(-0.0380)Y_{Knee}+(-0.026)Y_{A} \quad (2)$$

In the first canonical variable in the body_measurement set, all variables are negative contributions to the first canonical variable. When we increased neck variable by one unit, then V_1 decreased by 0.0337 units. Similar to how we can interpret the other variables, we can also write models for V_2, V_3, V_4 and V_5 same format

The number of canonical dimensions the same as the count of variables in the smaller set. The number of canonical dimensions that are significant in explaining the relationship between the 2 sets of variables may, however, be smaller than the number of variables in the smaller data set. In this case, there are 5 dimensions

3.4 Canonical Correlation and Squared correlation

```
[1] 0.9860315 0.7464733 0.6633413 0.2321466 0.1052086

[1] "Squared Canonical Correlations:"
[1] 0.97225818 0.55722241 0.44002167 0.05389205 0.01106886
```

$corr(U_1, V_1) = 0.9860$
$corr(U_2, V_2) = 0.7464$
$corr(U_3, V_3) = 0.6633$
$corr(U_4, V_4) = 0.2321$
$corr(U_5, V_5) = 0.1052$
This way to we can write the correlation between canonical variate pair

The Squared values of the Canonical variate pairs can be interpreted in the Same way as R^2 values are interpreted.97.2% of the variation of U_1 is explained by the Variation in V_1, and 55.7% variation of U_2 is explained by the Variation in V_2 similarly we can interpret other canonical variate pairs.These First one is highest canonical correlation and next one is moderate canonical correlation and this implies that First two canonical correlation are important

3.5 Canonical Loading

```
$corr.X.xscores
[,1]        [,2]        [,3]        [,4]        [,5]
Density  0.69532338  0.5381107  0.45434200  0.14277165 -0.01232537
BodyFat -0.71201857 -0.5403512 -0.42315842 -0.05588656  0.13734542
Age     -0.08006284 -0.8504807  0.45283366  0.19555338 -0.16423477
Weight  -0.98706445  0.1261738  0.07222592 -0.05585627  0.03804526
Height  -0.19189530  0.2274742  0.47502310 -0.71243498  0.42216242
```

```
$corr.Y.xscores
[,1]          [,2]          [,3]          [,4]          [,5]
Neck     -0.8236140  0.037796684  0.17708226 -0.008174081 -0.0068153945
Chest    -0.9200148 -0.103921142  0.01657949  0.038874478  0.0042745936
Abdomen  -0.9335578 -0.219475050 -0.07270940  0.009069461 -0.0009652415
Hip      -0.9422635  0.124190515 -0.05257039  0.020522120  0.0021816506
Thigh    -0.8622424  0.227620822 -0.16250724 -0.016069099 -0.0072390006
Knee     -0.8401756  0.095518583  0.10397375 -0.064914522  0.0176759705
Ankle    -0.5834644  0.209107655  0.09252320 -0.098098641 -0.0204341927
Biceps   -0.7928052  0.122436584  0.01020220 -0.046744114 -0.0244446613
Forearm  -0.6131744  0.142884729  0.01890613 -0.071676311  0.0436698902
Wrist    -0.7107346  0.008220391  0.36647237 -0.017782161 -0.0069705436
```

```
$corr.X.yscores
[,1]         [,2]         [,3]         [,4]         [,5]
Density  0.68561078  0.40168525  0.30138381  0.03314395 -0.001296736
BodyFat -0.70207276 -0.40335777 -0.28069845 -0.01297387  0.014449926
Age     -0.07894449 -0.63486115  0.30038327  0.04539705 -0.017278918
Weight  -0.97327667  0.09418537  0.04791044 -0.01296684  0.004002690
Height  -0.18921482  0.16980345  0.31510244 -0.16538937  0.044415138
```

```
$corr.Y.yscores
[,1]        [,2]        [,3]        [,4]         [,5]
Neck     -0.8352816  0.05063367  0.26695497 -0.03521086 -0.064779794
Chest    -0.9330481 -0.13921615  0.02499390  0.16745658  0.040629679
Abdomen  -0.9467829 -0.29401593 -0.10961084  0.03906781 -0.009174545
Hip      -0.9556119  0.16636966 -0.07925089  0.08840155  0.020736418
Thigh    -0.8744573  0.30492828 -0.24498284 -0.06921961 -0.068806137
Knee     -0.8520778  0.12795981  0.15674247 -0.27962727  0.168008722
Ankle    -0.5917299  0.28012743  0.13948054 -0.42257193 -0.194225409
Biceps   -0.8040364  0.16402004  0.01538002 -0.20135601 -0.232344601
Forearm  -0.6218608  0.19141304  0.02850136 -0.30875450  0.415078903
Wrist    -0.7208031  0.01101230  0.55246427 -0.07659884 -0.066254473
```

3.6 Hypothesis Testing

$H_0 : \rho_1^* = \rho_2^* = \rho_3^* = \rho_4^* = \rho_5^* = 0$
$H_1 : \rho_i^* \neq 0 \;\; \forall$i

```
Wilks' Lambda, using F-approximation (Rao's F):
                stat      approx df1       df2   p.value
1 to 5:  0.006435741 43.8775925  50 1084.2505 0.0000000
2 to 5:  0.231986971 11.8358970  36  893.6335 0.0000000
3 to 5:  0.523935667  7.2168734  24  693.7739 0.0000000
4 to 5:  0.935635615  1.1596814  14  480.0000 0.3033729
5 to 5:  0.988931140  0.4495755   6  241.0000 0.8448980

Hotelling-Lawley Trace, using F-approximation:
                stat       approx df1  df2   p.value
1 to 5:  37.15906941 174.9448988  50 1177 0.0000000
2 to 5:   2.11240822  13.9301586  36 1187 0.0000000
3 to 5:   0.85393799   8.5180314  24 1197 0.0000000
4 to 5:   0.06815459   1.1751798  14 1207 0.2882622
5 to 5:   0.01119275   0.4540526   6 1217 0.8423977

Pillai-Bartlett Trace, using F-approximation:
                stat      approx df1  df2   p.value
1 to 5:  2.03446317 16.5334526  50 1205 0.0000000
2 to 5:  1.06220499  9.1039321  36 1215 0.0000000
3 to 5:  0.50498258  5.7341608  24 1225 0.0000000
4 to 5:  0.06496091  1.1611823  14 1235 0.2996034
5 to 5:  0.01106886  0.4603768   6 1245 0.8378750
```

The above out put show the significant test.Here proceeding test until we found non significant results ,the first test determines whether the combined dimensions from 1 to 5 are significant. Since the p-value is less than the $\alpha = 0.05$ level of significance, it follows that all the 5 dimensions are statistically significant.

Similarly, the second test determines the significance of dimension from 2 to 5 combined. Since p<0.05, it follows that the dimensions are statistically significant.

Similary, the third test determine the significance of dimension from 3 to 5 combined.Since $p < 0.05$ it follows that the dimensions are statistically significant

The test 4 and 5 th test determine the significance of dimension from 4 to 5 and 5 to 5 combine.Since $p > 0.05$ it follows that the dimensions are statistically not significant.

Therefore First three pair is canonical variate pair is correlated.

3.7 Standardize the Data

Positive coefficients indicate that higher values of the variable contribute to higher values of the canonical variate.Negative coefficients indicate that higher values of the variable contribute to lower values of the canonical variate.Variables with larger absolute values of coefficients are more influential in the canonical variate.

```
               [,1]       [,2]       [,3]       [,4]       [,5]
Density  0.08136008  0.2389375  0.7772536  3.6877800  5.2571012
BodyFat -0.03301416 -0.4985470 -0.1977622  3.1724568  5.8299414
Age     -0.04611005 -0.6552077  0.7926119  0.1444547 -0.3159917
Weight  -0.94383605  0.6084133  0.5588144  0.4747515 -0.5844582
Height   0.08023630 -0.1405549  0.3450233 -0.9110396  0.5552816
```

$$U_1 = 0.0812 * X_{Density} + -0.0330 * X_{BodyFat} + -0.046 * X_{Age} + -0.943 * X_{Weight} + 0.080 * X_{Height} \quad (3)$$

We can write other 4 Canonical variate this Way.Here Density and Height variables are positive correlated others are Negatively Associated .In Weight variable has highest coefficient.The 0.0812 is meaning small positive relation ship to Canonical variate U_1

```
Neck    -0.08199892  0.24576999  0.49737184 -0.06238594 -0.1646555
Chest   -0.19237034  0.38205936  0.23342126  1.48661491  0.3629568
Abdomen -0.26177857 -2.16940847 -0.94206371 -1.33080104 -0.3590492
Hip     -0.30351227  1.14967040  0.34986024  1.40835387  0.5566160
Thigh   -0.03684225  0.58799402 -1.06672206  0.03356982 -0.4921687
Knee    -0.09179097 -0.14342100  0.59010606 -0.81024215  0.7733225
Ankle   -0.04546043  0.16213200 -0.11380535 -0.59644580 -0.4385774
Biceps  -0.06354585 -0.04815300  0.02227707 -0.48054635 -1.0376071
Forearm -0.01413291  0.07707143 -0.30814008 -0.41673542  1.0865164
Wrist   -0.02347024 -0.14368107  0.82601990  0.33110795 -0.2823361
```

$$V_1 = -0.0819*Y_{Neck}+-0.1923*Y_{Chest}+-0.2617*Y_{Abdomen}+-0.3035*Y_{Hip}+-0.0368*Y_{Thigh}+-0.0917*Y_{Knee}+-0.0454*Y_{Ank} \quad (4)$$

In First canonical variate all are negatively associated also all variables are small influential.

4 Conclusion

Canonical analysis is a method for exploring the relationship between two multivariate tests. Here we are exploring the relationship between body_consumption and body_measurement . In the body consumption set, body fat and density variables are highly associated with each other, and other variables have relatively small relationships with each other.In body measurement, most of the variables are highly correlated with each other.

The number of canonical dimensions is the same as the count of variables in the smaller set. That's why, in this case, we selected five dimensions. The positive coefficient indicates a highly significant association, and the negative coefficient indicates a small association with the corresponding canonical variates. The density variable has a high influence on all canonical variables. In body measurement variables, all are negative coefficients to fist coefficients.97.2% of the variation of U_1 is explained by the variation of V_1. $corr(U_1, V_1) = 0.98$; this correlated to a strong positive relationship (first canonical variate pairs). From Hyphothsis, the first three canonical variate pairs are significant.Abdomen and hip variables have a high influence on the second canonical varite pair.

consider i^{th} canonical variate pairs:

$$var(U_i) = var(V_i) = 1 \quad \forall \text{i} = 1{,}2{,}3{,}4{,}5$$
$$cov(U_i, U_j) = cov(V_i, V_j) = 0 \quad \forall \text{i} = 1{,}2{,}3{,}4{,}5 \quad i \neq j$$
$$cov(U_i, V_j) = 0 \quad i \neq j$$

These are the Limitation of the Canonical Correlation analysis.

References

[1] Amos Okutse. Canonical correlation analysis (cca) in r: A non-technical primer. *Analytics Vidhya*, January 2020. Accessed: 2024-05-26.

[2] Fede Soriano. Body fat prediction dataset. https://www.kaggle.com/datasets/fedesoriano/body-fat-prediction-dataset, 2024. Accessed: 2024-04-28.

[1] [2]

5 Appendices

5.1 CSV Appendix

Density	BodyFat	Age	Weight	Height	Neck	Chest	Abdomen	Hip	Thigh	Knee	Ankle	Biceps	Fore
1.0708	12.3	23	154.25	67.75	36.2	93.1	85.2	94.5	59.0	37.3	21.9	32.0	27.4
1.0853	6.1	22	173.25	72.25	38.5	93.6	83.0	98.7	58.7	37.3	23.4	30.5	28.9
1.0414	25.3	22	154.00	66.25	34.0	95.8	87.9	99.2	59.6	38.9	24.0	28.8	25.2
1.0751	10.4	26	184.75	72.25	37.4	101.8	86.4	101.2	60.1	37.3	22.8	32.4	29.4
1.0340	28.7	24	184.25	71.25	34.4	97.3	100.0	101.9	63.2	42.2	24.0	32.2	27.7
1.0502	20.9	24	210.25	74.75	39.0	104.5	94.4	107.8	66.0	42.0	25.6	35.7	30.6
1.0549	19.2	26	181.00	69.75	36.4	105.1	90.7	100.3	58.4	38.3	22.9	31.9	27.8
1.0704	12.4	25	176.00	72.50	37.8	99.6	88.5	97.1	60.0	39.4	23.2	30.5	29.0
1.0900	4.1	25	191.00	74.00	38.1	100.9	82.5	99.9	62.9	38.3	23.8	35.9	31.1
1.0722	11.7	23	198.25	73.50	42.1	99.6	88.6	104.1	63.1	41.7	25.0	35.6	30.0
1.0830	7.1	26	186.25	74.50	38.5	101.5	83.6	98.2	59.7	39.7	25.2	32.8	29.4
1.0812	7.8	27	216.00	76.00	39.4	103.6	90.9	107.7	66.2	39.2	25.9	37.2	30.2
1.0513	20.8	32	180.50	69.50	38.4	102.0	91.6	103.9	63.4	38.3	21.5	32.5	28.6
1.0505	21.2	30	205.25	71.25	39.4	104.1	101.8	108.6	66.0	41.5	23.7	36.9	31.6
1.0484	22.1	35	187.75	69.50	40.5	101.3	96.4	100.1	69.0	39.0	23.1	36.1	30.5
1.0512	20.9	35	162.75	66.00	36.4	99.1	92.8	99.2	63.1	38.7	21.7	31.1	26.4
1.0333	29.0	34	195.75	71.00	38.9	101.9	96.4	105.2	64.8	40.8	23.1	36.2	30.8
1.0468	22.9	32	209.25	71.00	42.1	107.6	97.5	107.0	66.9	40.0	24.4	38.2	31.6
1.0622	16.0	28	183.75	67.75	38.0	106.8	89.6	102.4	64.2	38.7	22.9	37.2	30.5
1.0610	16.5	33	211.75	73.50	40.0	106.2	100.5	109.0	65.8	40.6	24.0	37.1	30.1
1.0551	19.1	28	179.00	68.00	39.1	103.3	95.9	104.9	63.5	38.0	22.1	32.5	30.3
1.0640	15.2	28	200.50	69.75	41.3	111.4	98.8	104.8	63.4	40.6	24.6	33.0	32.8
1.0631	15.6	31	140.25	68.25	33.9	86.0	76.4	94.6	57.4	35.3	22.2	27.9	25.9
1.0584	17.7	32	148.75	70.00	35.5	86.7	80.0	93.4	54.9	36.2	22.1	29.8	26.7
1.0668	14.0	28	151.25	67.75	34.5	90.2	76.3	95.8	58.4	35.5	22.9	31.1	28.0
1.0911	3.7	27	159.25	71.50	35.7	89.6	79.7	96.5	55.0	36.7	22.5	29.9	28.2
1.0811	7.9	34	131.50	67.50	36.2	88.6	74.6	85.3	51.7	34.7	21.4	28.7	27.0
1.0468	22.9	31	148.00	67.50	38.8	97.4	88.7	94.7	57.5	36.0	21.0	29.2	26.6
1.0910	3.7	27	133.25	64.75	36.4	93.5	73.9	88.5	50.1	34.5	21.3	30.5	27.9
1.0790	8.8	29	160.75	69.00	36.7	97.4	83.5	98.7	58.9	35.3	22.6	30.1	26.7
1.0716	11.9	32	182.00	73.75	38.7	100.5	88.7	99.8	57.5	38.7	33.9	32.5	27.7
1.0862	5.7	29	160.25	71.25	37.3	93.5	84.5	100.6	58.5	38.8	21.5	30.1	26.4
1.0719	11.8	27	168.00	71.25	38.1	93.0	79.1	94.5	57.3	36.2	24.5	29.0	30.0
1.0502	21.3	41	218.50	71.00	39.8	111.7	100.5	108.3	67.1	44.2	25.2	37.5	31.5
1.0263	32.3	41	247.25	73.50	42.1	117.0	115.6	116.1	71.2	43.3	26.3	37.3	31.7
1.0101	40.1	49	191.75	65.00	38.4	118.5	113.1	113.8	61.9	38.3	21.9	32.0	29.8
1.0438	24.2	40	202.25	70.00	38.5	106.5	100.9	106.2	63.5	39.9	22.6	35.1	30.6
1.0346	28.4	50	196.75	68.25	42.1	105.6	98.8	104.8	66.0	41.5	24.7	33.2	30.5
1.0202	35.2	46	363.15	72.25	51.2	136.2	148.1	147.7	87.3	49.1	29.6	45.0	29.0
1.0258	32.6	50	203.00	67.00	40.2	114.8	108.1	102.5	61.3	41.1	24.7	34.1	31.0
1.0217	34.5	45	262.75	68.75	43.2	128.3	126.2	125.6	72.5	39.6	26.6	36.4	32.7
1.0250	32.9	44	205.00	29.50	36.6	106.0	104.3	115.5	70.6	42.5	23.7	33.6	28.7
1.0279	31.6	48	217.00	70.00	37.3	113.3	111.2	114.1	67.7	40.9	25.0	36.7	29.8
1.0269	32.0	41	212.00	71.50	41.5	106.6	104.3	106.0	65.0	40.2	23.0	35.8	31.5
1.0814	7.7	39	125.25	68.00	31.5	85.1	76.0	88.2	50.0	34.7	21.0	26.1	23.1
1.0670	13.9	43	164.25	73.25	35.7	96.6	81.5	97.2	58.4	38.2	23.4	29.7	27.4
1.0742	10.8	40	133.50	67.50	33.6	88.2	73.7	88.5	53.3	34.5	22.5	27.9	26.2
1.0665	5.6	39	148.50	71.25	34.6	89.8	79.5	92.7	52.7	37.5	21.9	28.8	26.8
1.0678	13.6	45	135.75	68.50	32.8	92.3	83.4	90.4	52.0	35.8	20.6	28.8	25.5
1.0903	4.0	47	127.50	66.75	34.0	83.4	70.4	87.2	50.6	34.4	21.9	26.8	25.8
1.0756	10.2	47	158.25	72.25	34.9	90.2	86.7	98.3	52.6	37.2	22.4	26.0	25.8
1.0840	6.6	40	139.25	69.00	34.3	89.2	77.9	91.0	51.4	34.9	21.0	26.7	26.1
1.0807	8.0	51	137.25	67.75	36.5	89.7	82.0	89.1	49.3	33.7	21.4	29.6	26.0
1.0848	6.3	49	152.75	73.50	35.1	93.3	79.6	91.6	52.6	37.6	22.6	38.5	27.4
1.0906	3.9	42	136.25	67.50	37.8	87.6	77.6	88.6	51.9	34.9	22.5	27.7	27.5
1.0473	22.6	54	198.00	72.00	39.9	107.6	100.0	99.6	57.2	38.0	22.0	35.9	30.2
1.0524	20.4	58	181.50	68.00	39.1	100.0	99.8	102.5	62.1	39.6	22.5	33.1	28.3
1.0356	28.0	62	201.25	69.50	40.5	111.5	104.2	105.8	61.8	39.8	22.7	37.7	30.9
1.0280	31.5	54	202.50	70.75	40.5	115.4	105.3	97.0	59.1	38.0	22.5	31.6	28.8
1.0430	24.6	61	179.75	65.75	38.4	104.8	98.3	99.6	60.6	37.7	22.9	34.5	29.6
1.0396	26.1	62	216.00	73.25	41.4	112.3	104.8	103.1	61.6	40.9	23.1	36.2	31.8

5.2 PDF Appendix

CanonicalCorrelation

S18809

2024-05-26

```r
library(ggplot2)
```

Loading Suitable Package

```
## Warning: package 'ggplot2' was built under R version 4.4.0
```

```r
library(GGally)
```

```
## Warning: package 'GGally' was built under R version 4.4.0

## Registered S3 method overwritten by 'GGally':
##   method from
##   +.gg   ggplot2
```

```r
library(CCA)#facilitates canonical correlation analysis
```

```
## Warning: package 'CCA' was built under R version 4.4.0

## Loading required package: fda

## Warning: package 'fda' was built under R version 4.4.0

## Loading required package: splines

## Loading required package: fds

## Warning: package 'fds' was built under R version 4.4.0

## Loading required package: rainbow

## Warning: package 'rainbow' was built under R version 4.4.0

## Loading required package: MASS

## Loading required package: pcaPP

## Warning: package 'pcaPP' was built under R version 4.4.0

## Loading required package: RCurl

## Warning: package 'RCurl' was built under R version 4.4.0

## Loading required package: deSolve

## Warning: package 'deSolve' was built under R version 4.4.0

##
## Attaching package: 'fda'

## The following object is masked from 'package:graphics':
##
##     matplot
```

```
## Loading required package: fields

## Warning: package 'fields' was built under R version 4.4.0

## Loading required package: spam

## Spam version 2.10-0 (2023-10-23) is loaded.
## Type 'help( Spam)' or 'demo( spam)' for a short introduction
## and overview of this package.
## Help for individual functions is also obtained by adding the
## suffix '.spam' to the function name, e.g. 'help( chol.spam)'.

##
## Attaching package: 'spam'

## The following objects are masked from 'package:base':
##
##     backsolve, forwardsolve

## Loading required package: viridisLite

##
## Try help(fields) to get started.
```

```r
library(CCP)#facilitates checking the significance of the canonical variates
```

```
## Warning: package 'CCP' was built under R version 4.4.0
```

```r
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.4.0

##
## Attaching package: 'dplyr'

## The following object is masked from 'package:MASS':
##
##     select

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(tidyverse)
```

```
## Warning: package 'readr' was built under R version 4.4.0

## Warning: package 'purrr' was built under R version 4.4.0

## Warning: package 'forcats' was built under R version 4.4.0

## Warning: package 'lubridate' was built under R version 4.4.0

## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v forcats   1.0.0     v stringr   1.5.1
## v lubridate 1.9.3     v tibble    3.2.1
## v purrr     1.0.2     v tidyr     1.3.1
## v readr     2.1.5
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x tidyr::complete() masks RCurl::complete()
```

```
## x dplyr::filter()   masks stats::filter()
## x dplyr::lag()      masks stats::lag()
## x dplyr::select()   masks MASS::select()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become error
```

```r
library(corrplot)
```

```
## Warning: package 'corrplot' was built under R version 4.4.0
```

```
## corrplot 0.92 loaded
```

```r
bodyFat <- read_csv(file = "../Data/bodyfat.csv")
```

Loading the File

```
## Rows: 252 Columns: 15
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl (15): Density, BodyFat, Age, Weight, Height, Neck, Chest, Abdomen, Hip, ...
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
colnames(bodyFat)
```

```
##  [1] "Density" "BodyFat" "Age"     "Weight"  "Height"  "Neck"    "Chest"
##  [8] "Abdomen" "Hip"     "Thigh"   "Knee"    "Ankle"   "Biceps"  "Forearm"
## [15] "Wrist"
```

```r
colSums(is.na(bodyFat))
```

Check the null values for all columns

```
## Density BodyFat     Age  Weight  Height    Neck   Chest Abdomen     Hip   Thigh
##       0       0       0       0       0       0       0       0       0       0
##    Knee   Ankle  Biceps Forearm   Wrist
##       0       0       0       0       0
```

```r
dim(bodyFat)
```

Dimension of the data

```
## [1] 252  15
```

```r
summary(bodyFat)
```

Summary of the Data set

```
##     Density         BodyFat           Age            Weight
##  Min.   :0.995   Min.   : 0.00   Min.   :22.00   Min.   :118.5
##  1st Qu.:1.041   1st Qu.:12.47   1st Qu.:35.75   1st Qu.:159.0
##  Median :1.055   Median :19.20   Median :43.00   Median :176.5
##  Mean   :1.056   Mean   :19.15   Mean   :44.88   Mean   :178.9
##  3rd Qu.:1.070   3rd Qu.:25.30   3rd Qu.:54.00   3rd Qu.:197.0
##  Max.   :1.109   Max.   :47.50   Max.   :81.00   Max.   :363.1
```

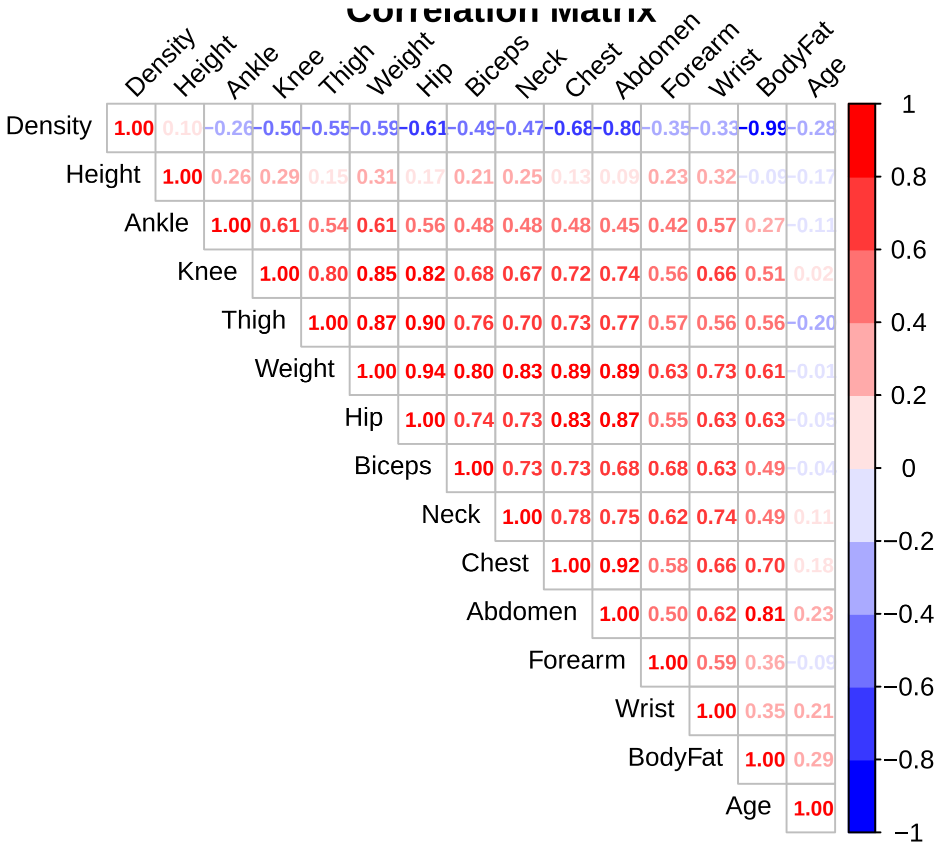
```
##      Height           Neck           Chest           Abdomen     
##  Min.   :29.50   Min.   :31.10   Min.   : 79.30   Min.   : 69.40  
##  1st Qu.:68.25   1st Qu.:36.40   1st Qu.: 94.35   1st Qu.: 84.58  
##  Median :70.00   Median :38.00   Median : 99.65   Median : 90.95  
##  Mean   :70.15   Mean   :37.99   Mean   :100.82   Mean   : 92.56  
##  3rd Qu.:72.25   3rd Qu.:39.42   3rd Qu.:105.38   3rd Qu.: 99.33  
##  Max.   :77.75   Max.   :51.20   Max.   :136.20   Max.   :148.10  
##       Hip            Thigh            Knee           Ankle           Biceps     
##  Min.   : 85.0   Min.   :47.20   Min.   :33.00   Min.   :19.1   Min.   :24.80  
##  1st Qu.: 95.5   1st Qu.:56.00   1st Qu.:36.98   1st Qu.:22.0   1st Qu.:30.20  
##  Median : 99.3   Median :59.00   Median :38.50   Median :22.8   Median :32.05  
##  Mean   : 99.9   Mean   :59.41   Mean   :38.59   Mean   :23.1   Mean   :32.27  
##  3rd Qu.:103.5   3rd Qu.:62.35   3rd Qu.:39.92   3rd Qu.:24.0   3rd Qu.:34.33  
##  Max.   :147.7   Max.   :87.30   Max.   :49.10   Max.   :33.9   Max.   :45.00  
##     Forearm          Wrist      
##  Min.   :21.00   Min.   :15.80  
##  1st Qu.:27.30   1st Qu.:17.60  
##  Median :28.70   Median :18.30  
##  Mean   :28.66   Mean   :18.23  
##  3rd Qu.:30.00   3rd Qu.:18.80  
##  Max.   :34.90   Max.   :21.40
```

Canonical Correlation Analysis

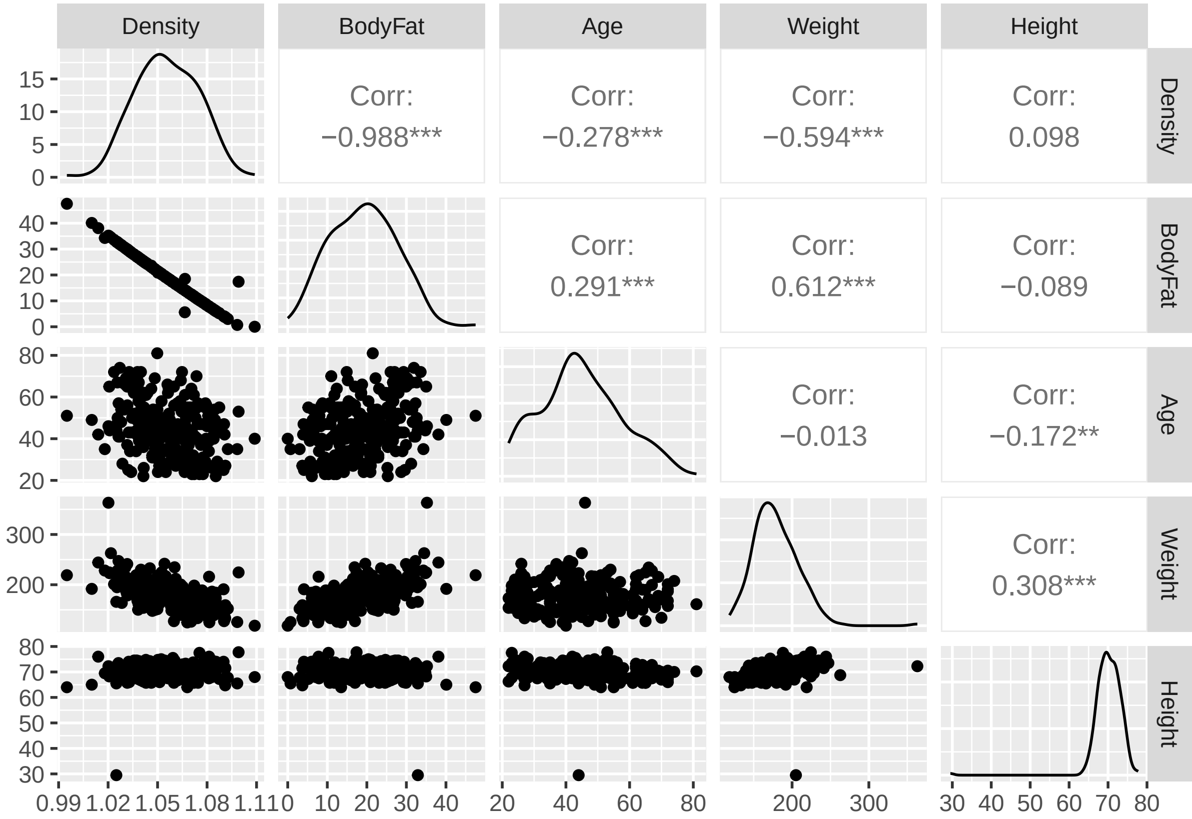
```r
body_composition_metrics <- bodyFat[, 1:5] #U_ Xvariables
body_measurements <- bodyFat[, 6:15] #V_ Y variables
```

Correlation matrix between each variables

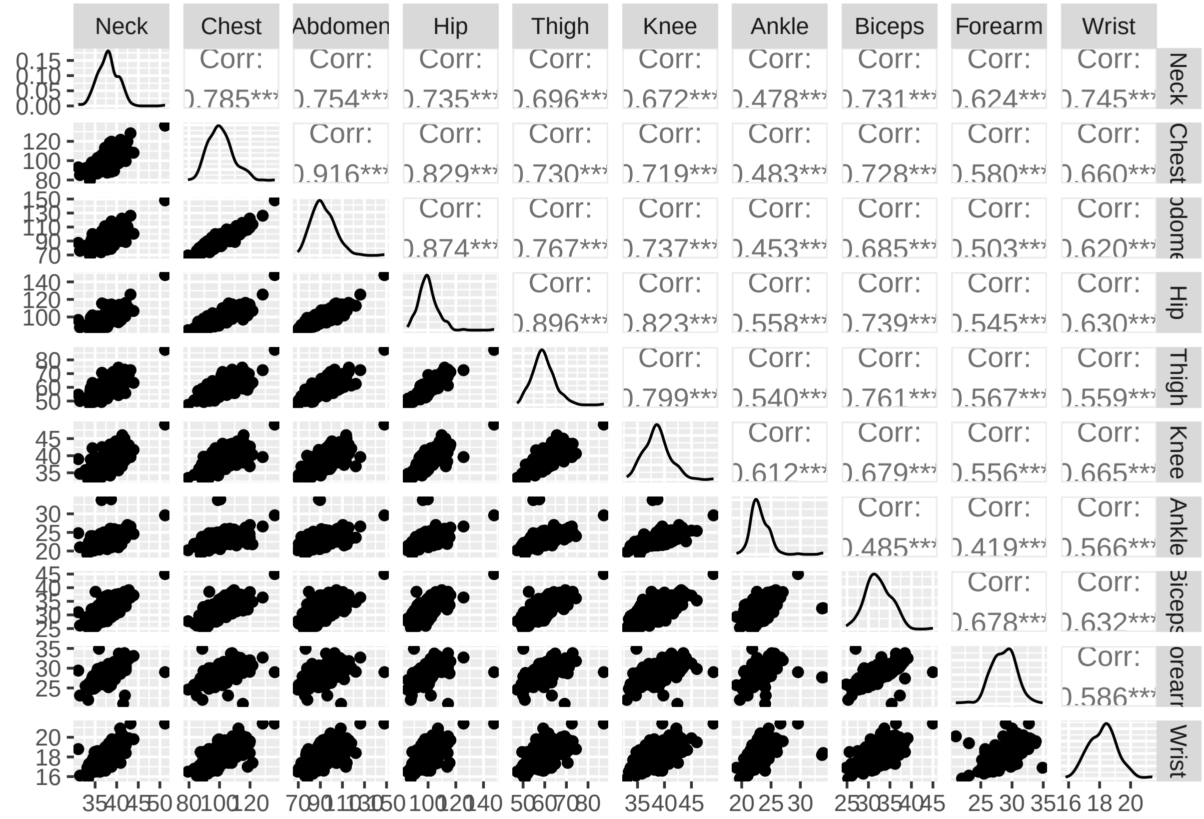
```r
# Compute correlation matrix
correlation_matrix <- cor(bodyFat)

# Draw correlation matrix with customization
corrplot(correlation_matrix, method = "number", col = colorRampPalette(c("blue", "white", "red"))(10),
         type = "upper", order = "hclust", tl.col = "black", tl.srt = 45,
         addrect = 3, rect.col = "grey", number.cex = 0.7, tl.cex = 0.8,
         title = "Correlation Matrix")
```




```r
ggpairs(body_composition_metrics)
```




```r
ggpairs(body_measurements)
```




```r
#checking the between and within set associations
matcor(body_composition_metrics,body_measurements)
```

```
## $Xcor
##             Density     BodyFat         Age      Weight      Height
## Density  1.00000000 -0.98778240 -0.27763721 -0.59406188  0.09788114
## BodyFat -0.98778240  1.00000000  0.29145844  0.61241400 -0.08949538
## Age     -0.27763721  0.29145844  1.00000000 -0.01274609 -0.17164514
## Weight  -0.59406188  0.61241400 -0.01274609  1.00000000  0.30827854
## Height   0.09788114 -0.08949538 -0.17164514  0.30827854  1.00000000
##
## $Ycor
##              Neck     Chest   Abdomen       Hip     Thigh      Knee     Ankle
## Neck    1.0000000 0.7848350 0.7540774 0.7349579 0.6956973 0.6724050 0.4778924
## Chest   0.7848350 1.0000000 0.9158277 0.8294199 0.7298586 0.7194964 0.4829879
## Abdomen 0.7540774 0.9158277 1.0000000 0.8740662 0.7666239 0.7371789 0.4532227
## Hip     0.7349579 0.8294199 0.8740662 1.0000000 0.8964098 0.8234726 0.5583868
## Thigh   0.6956973 0.7298586 0.7666239 0.8964098 1.0000000 0.7991703 0.5397971
## Knee    0.6724050 0.7194964 0.7371789 0.8234726 0.7991703 1.0000000 0.6116082
## Ankle   0.4778924 0.4829879 0.4532227 0.5583868 0.5397971 0.6116082 1.0000000
## Biceps  0.7311459 0.7279075 0.6849827 0.7392725 0.7614774 0.6787088 0.4848545
## Forearm 0.6236603 0.5801727 0.5033161 0.5450141 0.5668422 0.5558982 0.4190500
## Wrist   0.7448264 0.6601623 0.6198324 0.6300895 0.5586848 0.6645073 0.5661946
##            Biceps   Forearm     Wrist
## Neck    0.7311459 0.6236603 0.7448264
## Chest   0.7279075 0.5801727 0.6601623
```

```
## Abdomen 0.6849827 0.5033161 0.6198324
## Hip     0.7392725 0.5450141 0.6300895
## Thigh   0.7614774 0.5668422 0.5586848
## Knee    0.6787088 0.5558982 0.6645073
## Ankle   0.4848545 0.4190500 0.5661946
## Biceps  1.0000000 0.6782551 0.6321264
## Forearm 0.6782551 1.0000000 0.5855883
## Wrist   0.6321264 0.5855883 1.0000000
##
## $XYcor
##             Density     BodyFat         Age      Weight      Height       Neck
## Density  1.00000000 -0.98778240 -0.27763721 -0.59406188  0.09788114 -0.4729664
## BodyFat -0.98778240  1.00000000  0.29145844  0.61241400 -0.08949538  0.4905919
## Age     -0.27763721  0.29145844  1.00000000 -0.01274609 -0.17164514  0.1135052
## Weight  -0.59406188  0.61241400 -0.01274609  1.00000000  0.30827854  0.8307162
## Height   0.09788114 -0.08949538 -0.17164514  0.30827854  1.00000000  0.2537099
## Neck    -0.47296636  0.49059185  0.11350519  0.83071622  0.25370988  1.0000000
## Chest   -0.68259865  0.70262034  0.17644968  0.89419052  0.13489181  0.7848350
## Abdomen -0.79895463  0.81343228  0.23040942  0.88799494  0.08781291  0.7540774
## Hip     -0.60933143  0.62520092 -0.05033212  0.94088412  0.17039426  0.7349579
## Thigh   -0.55309098  0.55960753 -0.20009576  0.86869354  0.14843561  0.6956973
## Knee    -0.49504035  0.50866524  0.01751569  0.85316739  0.28605321  0.6724050
## Ankle   -0.26489003  0.26596977 -0.10505810  0.61368542  0.26474369  0.4778924
## Biceps  -0.48710872  0.49327113 -0.04116212  0.80041593  0.20781557  0.7311459
## Forearm -0.35164842  0.36138690 -0.08505555  0.63030143  0.22864922  0.6236603
## Wrist   -0.32571598  0.34657486  0.21353062  0.72977489  0.32206533  0.7448264
##              Chest     Abdomen         Hip      Thigh        Knee      Ankle
## Density -0.6825987 -0.79895463 -0.60933143 -0.5530910 -0.49504035 -0.2648900
## BodyFat  0.7026203  0.81343228  0.62520092  0.5596075  0.50866524  0.2659698
## Age      0.1764497  0.23040942 -0.05033212 -0.2000958  0.01751569 -0.1050581
## Weight   0.8941905  0.88799494  0.94088412  0.8686935  0.85316739  0.6136854
## Height   0.1348918  0.08781291  0.17039426  0.1484356  0.28605321  0.2647437
## Neck     0.7848350  0.75407737  0.73495788  0.6956973  0.67240498  0.4778924
## Chest    1.0000000  0.91582767  0.82941992  0.7298586  0.71949640  0.4829879
## Abdomen  0.9158277  1.00000000  0.87406618  0.7666239  0.73717888  0.4532227
## Hip      0.8294199  0.87406618  1.00000000  0.8964098  0.82347262  0.5583868
## Thigh    0.7298586  0.76662393  0.89640979  1.0000000  0.79917030  0.5397971
## Knee     0.7194964  0.73717888  0.82347262  0.7991703  1.00000000  0.6116082
## Ankle    0.4829879  0.45322269  0.55838682  0.5397971  0.61160820  1.0000000
## Biceps   0.7279075  0.68498272  0.73927252  0.7614774  0.67870883  0.4848545
## Forearm  0.5801727  0.50331609  0.54501412  0.5668422  0.55589819  0.4190500
## Wrist    0.6601623  0.61983243  0.63008954  0.5586848  0.66450729  0.5661946
##              Biceps     Forearm      Wrist
## Density -0.48710872 -0.35164842 -0.3257160
## BodyFat  0.49327113  0.36138690  0.3465749
## Age     -0.04116212 -0.08505555  0.2135306
## Weight   0.80041593  0.63030143  0.7297749
## Height   0.20781557  0.22864922  0.3220653
## Neck     0.73114592  0.62366027  0.7448264
## Chest    0.72790748  0.58017273  0.6601623
## Abdomen  0.68498272  0.50331609  0.6198324
## Hip      0.73927252  0.54501412  0.6300895
## Thigh    0.76147745  0.56684218  0.5586848
## Knee     0.67870883  0.55589819  0.6645073
```

```
## Ankle    0.48485454  0.41904999  0.5661946
## Biceps   1.00000000  0.67825513  0.6321264
## Forearm  0.67825513  1.00000000  0.5855883
## Wrist    0.63212642  0.58558825  1.0000000
```

```
#obtaining the canonical correlations
CC_bodyFat <- cc(body_composition_metrics,body_measurements)
```

Display the Canonical Correlation

```
CC_bodyFat$cor
```

```
## [1] 0.9860315 0.7464733 0.6633413 0.2321466 0.1052086
```

```
squared_canonical_correlations <- (CC_bodyFat$cor)^2
print("Squared Canonical Correlations:")
```

```
## [1] "Squared Canonical Correlations:"
```

```
print(squared_canonical_correlations)
```

```
## [1] 0.97225818 0.55722241 0.44002167 0.05389205 0.01106886
```

```
#raw canonical coefficients
CC_bodyFat[3:4]
```

```
## $xcoef
##                 [,1]        [,2]        [,3]         [,4]         [,5]
## Density  4.275036756 12.55488437 40.84051546 193.77309883 276.23252853
## BodyFat -0.003944937 -0.05957253 -0.02363106   0.37908414   0.69663308
## Age     -0.003658936 -0.05199219  0.06289553   0.01146281  -0.02507464
## Weight  -0.032115108  0.02070196  0.01901430   0.01615397  -0.01988686
## Height   0.021905393 -0.03837303  0.09419517  -0.24872386   0.15159801
##
## $ycoef
##                 [,1]        [,2]         [,3]        [,4]        [,5]
## Neck    -0.033731736  0.10110192  0.204602877 -0.02566358 -0.06773401
## Chest   -0.022818444  0.04531884  0.027687793  0.17633820  0.04305294
## Abdomen -0.024276798 -0.20118641 -0.087365019 -0.12341571 -0.03329748
## Hip     -0.042365972  0.16047755  0.048835486  0.19658606  0.07769563
## Thigh   -0.007017635  0.11199988 -0.203187010  0.00639431 -0.09374728
## Knee    -0.038059042 -0.05946626  0.244674075 -0.33594851  0.32064061
## Ankle   -0.026822001  0.09565911 -0.067146022 -0.35190756 -0.25876402
## Biceps  -0.021032800 -0.01593798  0.007373403 -0.15905422 -0.34343366
## Forearm -0.006994097  0.03814112 -0.152492420 -0.20623410  0.53769542
## Wrist   -0.025139906 -0.15390251  0.884782812  0.35466292 -0.30242144
```

```
# compute canonical loading
cc_bodyFat_2 <- comput(body_composition_metrics,body_measurements,CC_bodyFat)
```

```
#displays the canonical loading
cc_bodyFat_2[3:6]
```

```
## $corr.X.xscores
##                [,1]       [,2]        [,3]        [,4]        [,5]
## Density  0.69532338  0.5381107  0.45434200  0.14277165 -0.01232537
## BodyFat -0.71201857 -0.5403512 -0.42315842 -0.05588656  0.13734542
## Age     -0.08006284 -0.8504807  0.45283366  0.19555338 -0.16423477
```

```
## Weight  -0.98706445  0.1261738  0.07222592 -0.05585627  0.03804526
## Height  -0.19189530  0.2274742  0.47502310 -0.71243498  0.42216242
##
## $corr.Y.xscores
##               [,1]         [,2]        [,3]         [,4]          [,5]
## Neck    -0.8236140  0.037796684  0.17708226 -0.008174081 -0.0068153945
## Chest   -0.9200148 -0.103921142  0.01657949  0.038874478  0.0042745936
## Abdomen -0.9335578 -0.219475050 -0.07270940  0.009069461 -0.0009652415
## Hip     -0.9422635  0.124190515 -0.05257039  0.020522120  0.0021816506
## Thigh   -0.8622424  0.227620822 -0.16250724 -0.016069099 -0.0072390006
## Knee    -0.8401756  0.095518583  0.10397375 -0.064914522  0.0176759705
## Ankle   -0.5834644  0.209107655  0.09252320 -0.098098641 -0.0204341927
## Biceps  -0.7928052  0.122436584  0.01020220 -0.046744114 -0.0244446613
## Forearm -0.6131744  0.142884729  0.01890613 -0.071676311  0.0436698902
## Wrist   -0.7107346  0.008220391  0.36647237 -0.017782161 -0.0069705436
##
## $corr.X.yscores
##                [,1]        [,2]        [,3]        [,4]         [,5]
## Density  0.68561078  0.40168525  0.30138381  0.03314395 -0.001296736
## BodyFat -0.70207276 -0.40335777 -0.28069845 -0.01297387  0.014449926
## Age     -0.07894449 -0.63486115  0.30038327  0.04539705 -0.017278918
## Weight  -0.97327667  0.09418537  0.04791044 -0.01296684  0.004002690
## Height  -0.18921482  0.16980345  0.31510244 -0.16538937  0.044415138
##
## $corr.Y.yscores
##               [,1]        [,2]        [,3]        [,4]         [,5]
## Neck    -0.8352816  0.05063367  0.26695497 -0.03521086 -0.064779794
## Chest   -0.9330481 -0.13921615  0.02499390  0.16745658  0.040629679
## Abdomen -0.9467829 -0.29401593 -0.10961084  0.03906781 -0.009174545
## Hip     -0.9556119  0.16636966 -0.07925089  0.08840155  0.020736418
## Thigh   -0.8744573  0.30492828 -0.24498284 -0.06921961 -0.068806137
## Knee    -0.8520778  0.12795981  0.15674247 -0.27962727  0.168008722
## Ankle   -0.5917299  0.28012743  0.13948054 -0.42257193 -0.194225409
## Biceps  -0.8040364  0.16402004  0.01538002 -0.20135601 -0.232344601
## Forearm -0.6218608  0.19141304  0.02850136 -0.30875450  0.415078903
## Wrist   -0.7208031  0.01101230  0.55246427 -0.07659884 -0.066254473
```

Test the Canonical correlation

```
rho <- CC_bodyFat$cor
```

Define number of observations, number of variables in first set, and number of variables in the second set.

```
n <- dim(body_composition_metrics)[1]
p <- length(body_composition_metrics)
q <- length(body_measurements)
```

Calculate p-values using the F-approximations of different test statistics:

```
p.asym(rho, n, p, q, tstat = "Wilks")
```

```
## Wilks' Lambda, using F-approximation (Rao's F):
##               stat     approx df1       df2   p.value
## 1 to 5:  0.006435741 43.8775925  50 1084.2505 0.0000000
## 2 to 5:  0.231986971 11.8358970  36  893.6335 0.0000000
## 3 to 5:  0.523935667  7.2168734  24  693.7739 0.0000000
## 4 to 5:  0.935635615  1.1596814  14  480.0000 0.3033729
## 5 to 5:  0.988931140  0.4495755   6  241.0000 0.8448980
```

```r
p.asym(rho, n, p, q, tstat = "Hotelling")
```

```
##  Hotelling-Lawley Trace, using F-approximation:
##               stat      approx df1  df2   p.value
## 1 to 5:  37.15906941 174.9448988  50 1177 0.0000000
## 2 to 5:   2.11240822  13.9301586  36 1187 0.0000000
## 3 to 5:   0.85393799   8.5180314  24 1197 0.0000000
## 4 to 5:   0.06815459   1.1751798  14 1207 0.2882622
## 5 to 5:   0.01119275   0.4540526   6 1217 0.8423977
```

```r
p.asym(rho, n, p, q, tstat = "Pillai")
```

```
##  Pillai-Bartlett Trace, using F-approximation:
##              stat     approx df1  df2   p.value
## 1 to 5:  2.03446317 16.5334526  50 1205 0.0000000
## 2 to 5:  1.06220499  9.1039321  36 1215 0.0000000
## 3 to 5:  0.50498258  5.7341608  24 1225 0.0000000
## 4 to 5:  0.06496091  1.1611823  14 1235 0.2996034
## 5 to 5:  0.01106886  0.4603768   6 1245 0.8378750
```

```r
p.asym(rho, n, p, q, tstat = "Roy")
```

```
##  Roy's Largest Root, using F-approximation:
##             stat   approx df1 df2 p.value
## 1 to 1:  0.9722582 844.6245  10 241       0
##
##  F statistic for Roy's Greatest Root is an upper bound.
```

Standardize the data set

```r
# Standardize first canonical variables
std_1<-diag(sqrt(diag(cov(body_composition_metrics))))
ss_1<- std_1%*%CC_bodyFat$xcoef
`rownames<-`(ss_1,c("Density","BodyFat","Age","Weight","Height"))
```

```
##                [,1]       [,2]       [,3]       [,4]       [,5]
## Density  0.08136008  0.2389375  0.7772536  3.6877800  5.2571012
## BodyFat -0.03301416 -0.4985470 -0.1977622  3.1724568  5.8299414
## Age     -0.04611005 -0.6552077  0.7926119  0.1444547 -0.3159917
## Weight  -0.94383605  0.6084133  0.5588144  0.4747515 -0.5844582
## Height   0.08023630 -0.1405549  0.3450233 -0.9110396  0.5552816
```

```r
# Standardize second canonical variables
std_2<-diag(sqrt(diag(cov(body_measurements))))
ss_2<- std_2%*%CC_bodyFat$ycoef
`rownames<-`(ss_2,c("Neck","Chest","Abdomen","Hip","Thigh","Knee","Ankle","Biceps"
                    ,"Forearm","Wrist"))
```

```
##                [,1]        [,2]        [,3]        [,4]       [,5]
```

```
## Neck    -0.08199892  0.24576999  0.49737184 -0.06238594 -0.1646555
## Chest   -0.19237034  0.38205936  0.23342126  1.48661491  0.3629568
## Abdomen -0.26177857 -2.16940847 -0.94206371 -1.33080104 -0.3590492
## Hip     -0.30351227  1.14967040  0.34986024  1.40835387  0.5566160
## Thigh   -0.03684225  0.58799402 -1.06672206  0.03356982 -0.4921687
## Knee    -0.09179097 -0.14342100  0.59010606 -0.81024215  0.7733225
## Ankle   -0.04546043  0.16213200 -0.11380535 -0.59644580 -0.4385774
## Biceps  -0.06354585 -0.04815300  0.02227707 -0.48054635 -1.0376071
## Forearm -0.01413291  0.07707143 -0.30814008 -0.41673542  1.0865164
## Wrist   -0.02347024 -0.14368107  0.82601990  0.33110795 -0.2823361
```