Exercise 7.6

02409 Multivariate Statistics

`Week E`

Different animal breeds:

- Hypertrophied Piemontese n=23
- Normal Piemontese n=12
- Hypertrophied Piemontese x Friesian Crossbreed n=10
- Friesian n=11
- Belgian blue and white n=23

18 different measurements were taken on each of the 79 young bulls. The chemical measurements are:

1. pH (pH-value)
2. Water
3. Protein
4. EtherExt (ether extract - related to fat content)
5. Hydroxy (hydroxyproline content, i.e. cartilage)
6. CollaSol (heat-solubility of collagen)

The physical parameters are:

7. Lightn (lightness)
8. Hue (relates to the color of the meat)
9. Driploss
10. CookLoss (cooking loss)
11. WBshear (Warner-Bratzler shear, which gives an indication of the tenderness of the meat)

And finally seven different sensory parameters (A, Te, Tf, Tr, Ji, Js, Oa) were assesed using a trained panel.

12. Appear (appearance)
13. EaseSink (ease of sinking)
14. Friabil (friability)
15. Residue (after chewing)
16. InJuice (initial juiciness)
17. SusJuice (sustained juiciness)
18. OvAcc (overall acceptability

```
Beef <- data.frame("pH" = c(1,rep(0,17)),
                "Water" = c(0.09,1,rep(0,16)),
                "Protein"  = c(0.28,-0.4,1,rep(0,15)),
                "EtherExt" = c(-0.28,-0.16,-0.56,1,rep(0,14)),
                "Hydroxy" = c(-0.33,-0.08,-0.55,0.59,1,rep(0,13)),
                "CollaSol" = c(-0.08,-0.01,-0.03,0.05,0.16,1,rep(0,12)),
                "Lightn" = c(-0.02,0.03,0.34,-0.31,-0.48,-0.02,1,rep(0,11)),
                "Hue" = c(-0.33,-0.23,-0.47,0.4,0.62,-0.03,-0.21,1,rep(0,10)),
                "DripLoss" = c(0.01,0.18,-0.07,0.08,-0.12,-0.1,0.25,-0.13,1,rep(0,
9)),
                "CookLoss" = c(-0.38,0.15,-0.64,0.44,0.66,-0.01,-0.45,0.65,0.03,1,
rep(0,8)),
                "WBshear" = c(-0.26,-0.01,-0.63,0.42,0.72,-0.03,-0.55,0.67,-0.11,
0.73,1,rep(0,7)),
                "Appear" = c(0.1,-0.003,0.25,-0.42,-0.33,-0.19,0.35,0.07,0.02,-0.1
8,-0.28,1,rep(0,6)),
                "EaseSink" = c(0.17,-0.16,0.27,-0.11,-0.26,0.01,0.19,-0.19,-0.02,-
0.36,-0.38,0.24,1,rep(0,5)),
                "Friabil" = c(0.1,-0.17,0.2,-0.09,-0.22,0.06,0.19,-0.1,-0.03,-0.3
1,-0.32,0.27,0.93,1,rep(0,4)),
                "Residue"  = c(0.08,-0.19,0.23,-0.13,-0.24,-0.02,0.2,-0.1,-0.02,-
0.32,-0.33,0.33,0.91,0.94,1,rep(0,3)),
                "InJuice" = c(0.08,-0.08,0.03,-0.004,-0.05,0.05,-0.06,-0.03,-0.13,
-0.12,-0.1,0.16,0.69,0.72,0.72,1,rep(0,2)),
                "SusJuice" = c(0.01,-0.09,-0.004,-0.01,-0.02,0.03,-0.02,0.08,-0.1
5,-0.07,-0.03,0.24,0.66,0.7,0.7,0.93,1,rep(0,1)),
                "OvAcc" = c(0.13,-0.13,0.21,-0.09,-0.22,0.07,0.22,-0.14,-0.01,-0.3
4,-0.37,0.31,0.92,0.92,0.91,0.8,0.79,1)
)
Beef = Beef + t(Beef) - diag(1,18,18)
```

```
# Recommended Packages:
library(CCP)
library(geigen)
```

Canonical Correlation Analysis on Beef Characterization

We consider the data from Exercise 7.3 and do now want to analyze the relation between the panel assessments and the physical/chemical measurements on the meat by means of Canonical Correlation Analyses of the data. A sample program for this is:

```
Exx = as.matrix(Beef[1:11,1:11])
Eyx = as.matrix(Beef[12:18,1:11])
Exy = as.matrix(Beef[1:11,12:18])
Eyy = as.matrix(Beef[12:18,12:18])
invExx = solve(Exx)
invEyy = solve(Eyy)

#Calculating the Canonical correlations:
Cancorr = geigen(Eyx%*%invExx%*%Exy,Eyy,symmetric = TRUE)
values = sort(Cancorr$values,decreasing = TRUE)

# E is the residual variation after having predicted Y by means of X
H = Eyx%*%invExx%*%Exy
E = Eyy - Eyx%*%invExx%*%Exy
invE = solve(E)
Ev <- eigen(invE%*%H)
var = Ev$values

# Eigenvalues, Proportion and Cumulative proportion of Variance:
varPC <- var/sum(var)
cumu = c(1:7)
for (i in 1:7){
  cumu[i] = sum(varPC[1:i])
}
results <- data.frame("CanCor" = sqrt(values),"Squared CanCor" = values,"eigenvalu
es"=var,"proportion"=varPC,
                      "cumulative" = cumu)
print("Table with information about the Canonical Correlations:")
results
```

```
## [1] "Table with information about the Canonical Correlations:"
```

```
##      CanCor Squared.CanCor eigenvalues  proportion cumulative
## 1 0.6720912     0.45170663  0.82384114 0.447388814  0.4473888
## 2 0.6148370     0.37802455  0.60778051 0.330056594  0.7774454
## 3 0.4036625     0.16294345  0.19466241 0.105711869  0.8831573
## 4 0.2987878     0.08927414  0.09802526 0.053232840  0.9363901
## 5 0.2636168     0.06949381  0.07468388 0.040557253  0.9769474
## 6 0.1637651     0.02681901  0.02755809 0.014965481  0.9919129
## 7 0.1211342     0.01467351  0.01489202 0.008087147  1.0000000
```

```
## Calculating p-values using F-approximations with Wilks test:
n = 79
p = length(Exx[,1])
q = length(Eyy[,1])


HypTest <- p.asym(results$CanCor,n,p,q,tstat="Wilks")

results_all <- data.frame(results,HypTest)
print("Table with information about the Canonical Correlations similar to the outp
ut in SAS:")
results_all
```

```
## Wilks' Lambda, using F-approximation (Rao's F):
##               stat    approx df1      df2    p.value
## 1 to 7:  0.2319646 1.3375749  77 372.9807 0.04166877
## 2 to 7:  0.4230666 0.9810705  60 329.8928 0.51995711
## 3 to 7:  0.6801983 0.5696448  45 284.9171 0.98797220
## 4 to 7:  0.8126073 0.4297982  32 237.6155 0.99724136
## 5 to 7:  0.8922633 0.3609966  21 187.1950 0.99592163
## 6 to 7:  0.9589010 0.2332596  12 132.0000 0.99637196
## 7 to 7:  0.9853265 0.1995531   5  67.0000 0.96156219
```

```
## [1] "Table with information about the Canonical Correlations similar to the out
put in SAS:"
```

```
##      CanCor Squared.CanCor eigenvalues  proportion cumulative    id      stat
## 1 0.6720912     0.45170663  0.82384114 0.447388814  0.4473888 Wilks 0.2319646
## 2 0.6148370     0.37802455  0.60778051 0.330056594  0.7774454 Wilks 0.4230666
## 3 0.4036625     0.16294345  0.19466241 0.105711869  0.8831573 Wilks 0.6801983
## 4 0.2987878     0.08927414  0.09802526 0.053232840  0.9363901 Wilks 0.8126073
## 5 0.2636168     0.06949381  0.07468388 0.040557253  0.9769474 Wilks 0.8922633
## 6 0.1637651     0.02681901  0.02755809 0.014965481  0.9919129 Wilks 0.9589010
## 7 0.1211342     0.01467351  0.01489202 0.008087147  1.0000000 Wilks 0.9853265
##      approx df1      df2    p.value
## 1 1.3375749  77 372.9807 0.04166877
## 2 0.9810705  60 329.8928 0.51995711
## 3 0.5696448  45 284.9171 0.98797220
## 4 0.4297982  32 237.6155 0.99724136
## 5 0.3609966  21 187.1950 0.99592163
## 6 0.2332596  12 132.0000 0.99637196
## 7 0.1995531   5  67.0000 0.96156219
```

1. How many Canonical Correlation Coefficients are significantly different from 0?

2. Give a verbal description of the definition of the first 2 canonical variables for the panel data and for the physical/chemical measurements!

3. What are the plots showing?

4. Describe the relationship between the panel assessments and the physical/chemical measurements!