Multivariate Analysis for the Behavioral Sciences, Second Edition (Chapman and Hall/CRC, 2019) **Solutions to Exercises of Chapter 18: Grouped Multivariate Data**

Kimmo Vehkalahti and Brian S. Everitt

20 December 2018

Solutions

Exercise 18.4

```r
pottery <- read.csv("data/pottery.csv")
str(pottery)
```

```
## 'data.frame':    45 obs. of  10 variables:
##  $ kiln : int  1 1 1 1 1 1 1 1 1 1 ...
##  $ AL2O3: num  18.8 16.9 18.2 16.9 17.8 18.8 16.5 18 15.8 14.6 ...
##  $ FE2O3: num  9.52 7.33 7.64 7.29 7.24 7.45 7.05 7.42 7.15 6.87 ...
##  $ MGO  : num  2 1.65 1.82 1.56 1.83 2.06 1.81 2.06 1.62 1.67 ...
##  $ CAO  : num  0.79 0.84 0.77 0.76 0.92 0.87 1.73 1 0.71 0.76 ...
##  $ NA2O : num  0.4 0.4 0.4 0.4 0.43 0.25 0.33 0.28 0.38 0.33 ...
##  $ K2O  : num  3.2 3.05 3.07 3.05 3.12 3.26 3.2 3.37 3.25 3.06 ...
##  $ TIO2 : num  1.01 0.99 0.98 1 0.93 0.98 0.95 0.96 0.93 0.91 ...
##  $ MNO  : num  0.077 0.067 0.087 0.063 0.061 0.072 0.066 0.072 0.062 0.055 ...
##  $ BAO  : num  0.015 0.018 0.014 0.019 0.019 0.017 0.019 0.017 0.017 0.012 ...
```

```r
head(pottery)
```

```
##   kiln AL2O3 FE2O3  MGO  CAO NA2O  K2O TIO2   MNO   BAO
## 1    1  18.8  9.52 2.00 0.79 0.40 3.20 1.01 0.077 0.015
## 2    1  16.9  7.33 1.65 0.84 0.40 3.05 0.99 0.067 0.018
## 3    1  18.2  7.64 1.82 0.77 0.40 3.07 0.98 0.087 0.014
## 4    1  16.9  7.29 1.56 0.76 0.40 3.05 1.00 0.063 0.019
## 5    1  17.8  7.24 1.83 0.92 0.43 3.12 0.93 0.061 0.019
## 6    1  18.8  7.45 2.06 0.87 0.25 3.26 0.98 0.072 0.017
```

```r
pottery_manova <- manova(cbind(AL2O3, FE2O3, MGO, CAO, NA2O, K2O, TIO2, MNO) ~ kiln,
                         data = pottery)

summary(pottery_manova, test = "Pillai")
```

```
##           Df  Pillai approx F num Df den Df    Pr(>F)
## kiln       1 0.92986   59.657      8     36 < 2.2e-16 ***
## Residuals 43
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
summary(pottery_manova, test = "Wilks")
```

```
##           Df   Wilks approx F num Df den Df    Pr(>F)
## kiln       1 0.07014   59.657      8     36 < 2.2e-16 ***
## Residuals 43
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
summary(pottery_manova, test = "Hotelling")
```

```
##           Df Hotelling-Lawley approx F num Df den Df    Pr(>F)
## kiln       1           13.257   59.657      8     36 < 2.2e-16 ***
## Residuals 43
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
summary(pottery_manova, test = "Roy")
```

```
##           Df    Roy approx F num Df den Df    Pr(>F)
## kiln       1 13.257   59.657      8     36 < 2.2e-16 ***
## Residuals 43
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
# all highly significant

region <- rep(1, length(pottery$kiln))
region[pottery$kiln == 2 | pottery$kiln == 3] <- 2
region[pottery$kiln == 4 | pottery$kiln == 5] <- 3

# calculate means of each region
m1 <- apply(pottery[region == 1, -1], 2, mean)
m2 <- apply(pottery[region == 2, -1], 2, mean)
m3 <- apply(pottery[region == 3, -1], 2, mean)

# find numbers in each class
n1 <- length(pottery[region==1, 1])
n2 <- length(pottery[region==2, 1])
n3 <- length(pottery[region==3, 1])

# find pooled covariance matrix
S123 <- ((n1-1)*var(pottery[region==1, -1]) +
         (n2-1)*var(pottery[region==2, -1]) +
         (n3-1)*var(pottery[region==3, -1])) / ((n1-1)+(n2-1)+(n3-1))

# find coefficients for each classification class
invS <- solve(S123)
a1 <- invS %*% (m1-m2)
a2 <- invS %*% (m1-m3)
a3 <- invS %*% (m2-m3)

# find thresholds
z12 <-(m1 %*% a1 + m2 %*% a1) / 2
z13 <-(m1 %*% a2 + m3 %*% a2) / 2
z23 <-(m2 %*% a3 + m3 %*% a3) / 2
```

```r
# new pot with the following chemical composition:
newvalues<-c(15.5, 5.71, 2.07, 0.98, 0.65, 3.01, 0.76, 0.09, 0.012)

(newvalues - as.vector(z12)) %*% a1
```

```
##           [,1]
## [1,] -2276.902
```

```r
(newvalues - as.vector(z13)) %*% a2
```

```
##           [,1]
## [1,] -10225.13
```

```r
(newvalues - as.vector(z23)) %*% a3
```

```
##          [,1]
## [1,] 29623.82
```

```r
# conclusion: allocate to region 2 (plot a graph!)
```