Classifying spectrometric curves

- **Statistical aim**

  We recall that for each unit i (among 215 pieces of finely chopped meat), we observe one spectrometric curve $(\boldsymbol{x}_i)$ which corresponds to the absorbance measured at 100 wavelengths (i.e. $\boldsymbol{x}_i = (\chi_i(\lambda_1), \ldots, \chi_i(\lambda_{100}))$. The first 100 columns of the file "npfda-spectrometric.dat" contain the observed curves $\{\boldsymbol{x}_i\}_{i=1,\ldots,215}$. It is worth noting that the reponse (fat content) is completely ignored for this classification analysis.

- **R/S+ commandlines**:

  - **Entering spectrometric data**

    ```
    SPECDAT <- as.matrix(read.table("npfda-spectrometric.dat"))
    attributes(SPECDAT)$dimnames[[1]] <- character(0)
    SPECURVES <- SPECDAT[,1:100]           # sample of curves
    ```

  - **Automatical classification: the following commandline allows to get directly the groups**:

    ```
    spec.classif <- classif.npfda(SPECURVES, 2, 20, c(0,1),
                    semimetric="deriv", threshold=0.05, nb.bw=100,
                    nss=0, mspg=10, centrality="mean")
    ```

    Remark: `nss=0` means that the criterion HI is used instead of SHI whereas `centrality="mean"` implies that HI is based on differences between mean and modal curves.

  - **Loading the groups**

    ```
    Group1 <- spec.classif$Partition[[1]]
    Group21 <- spec.classif$Partition[[2]]
    Group22 <- spec.classif$Partition[[3]]
    ```

    Remark: `spec.classif$Labels` contains the history of the splitting procedure namely '1', '21' and '22'.

- **Plotting terminal leaves: Group1, Group21 and Group22**
  - **Computing the second derivatives of the spectrometric curves**

```
SPECD2 <- t(approx.spline.deriv(SPECURVES, 2, 20,
            c(0,1))$APPROX)
```

  - **Loading the modal curve for each group**

```
MODAL.CURVES <- spec.classif$MODES
```

  - **Computing the second derivatives of the modal curves**

```
MODAL.CURVESD2 <- t(approx.spline.deriv(
                    MODAL.CURVES, 2, 20, c(0,1))$APPROX)
```

  - **Displaying groups: original data, second derivatives and corresponding modal curves**

    The following R/S+ commandlines allow to build Figure 1:

```
x <- seq(850, 1050, length=100)
##
# GROUP1
########
par(mfrow=c(3,3))
plot(x,SPECURVES[Group1[1],],ylim=range(SPECURVES), type="l",
     xlab="Wavelenghts",ylab="Absorbances", main="GROUP 1")
for(i in Group1[-1]){
        par(new=T)
        plot(x,SPECURVES[i,],ylim=range(SPECURVES), type="l",
             xlab="", ylab="", axes=F)
}
plot(x,SPECD2[Group1[1],],ylim=range(SPECD2), type="l",
     xlab="Wavelenghts", ylab="Absorbances",
     main="GROUP 1: second derivatives")
for(i in Group1[-1]){
        par(new=T)
        plot(x,SPECD2[i,],ylim=range(SPECD2), type="l",
             xlab="",ylab="", axes=F)
}
plot(x, MODAL.CURVESD2[1,], ylim=range(SPECD2), type="l",
     xlab="Wavelenghts", ylab="Absorbances",
     main="GROUP 1: modal curve")
##
```

```
# GROUP21
#########
plot(x,SPECURVES[Group21[1],],ylim=range(SPECURVES), type="l",
     xlab="Wavelenghts",ylab="Absorbances", main="GROUP 21")
for(i in Group21[-1]){
        par(new=T)
        plot(x,SPECURVES[i,],ylim=range(SPECURVES), type="l",
             xlab="", ylab="", axes=F)
}
plot(x,SPECD2[Group21[1],],ylim=range(SPECD2), type="l",
     xlab="Wavelenghts", ylab="Absorbances",
     main="GROUP 21: second derivatives")
for(i in Group21[-1]){
        par(new=T)
        plot(x,SPECD2[i,],ylim=range(SPECD2), type="l", xlab="",
             ylab="", axes=F)
}
plot(x, MODAL.CURVESD2[2,], ylim=range(SPECD2), type="l",
     xlab="Wavelenghts", ylab="Absorbances",
     main="GROUP 21: modal curve")
##
# GROUP22
#########
plot(x,SPECURVES[Group22[1],],ylim=range(SPECURVES), type="l",
     xlab="Wavelenghts",ylab="Absorbances", main="GROUP 22")
for(i in Group22[-1]){
        par(new=T)
        plot(x,SPECURVES[i,],ylim=range(SPECURVES), type="l",
             xlab="", ylab="", axes=F)
}
plot(x,SPECD2[Group22[1],],ylim=range(SPECD2), type="l",
     xlab="Wavelenghts", ylab="Absorbances",
     main="GROUP 22: second derivatives")
for(i in Group22[-1]){
        par(new=T)
        plot(x,SPECD2[i,],ylim=range(SPECD2), type="l", xlab="",
             ylab="", axes=F)
}
```

```
plot(x, MODAL.CURVESD2[3,], ylim=range(SPECD2), type="l",
     xlab="Wavelenghts", ylab="Absorbances",
     main="GROUP 22: modal curve")
```

•• Displaying the splitting score behavior along the procedure
The following commandlines allow to perform Figure 2:

```
Splitting.score <- spec.classif$Ssc
Split.names <- c(paste("GROUPS","\n","1 & 2"),
                 paste("GROUPS", "\n","21 & 22"))
par(mfrow=c(1,1))
barplot(Splitting.score, names=Split.names, main="Splitting scores",
        ylab="Percentages")
abline(h=0.05, lwd=3)
```

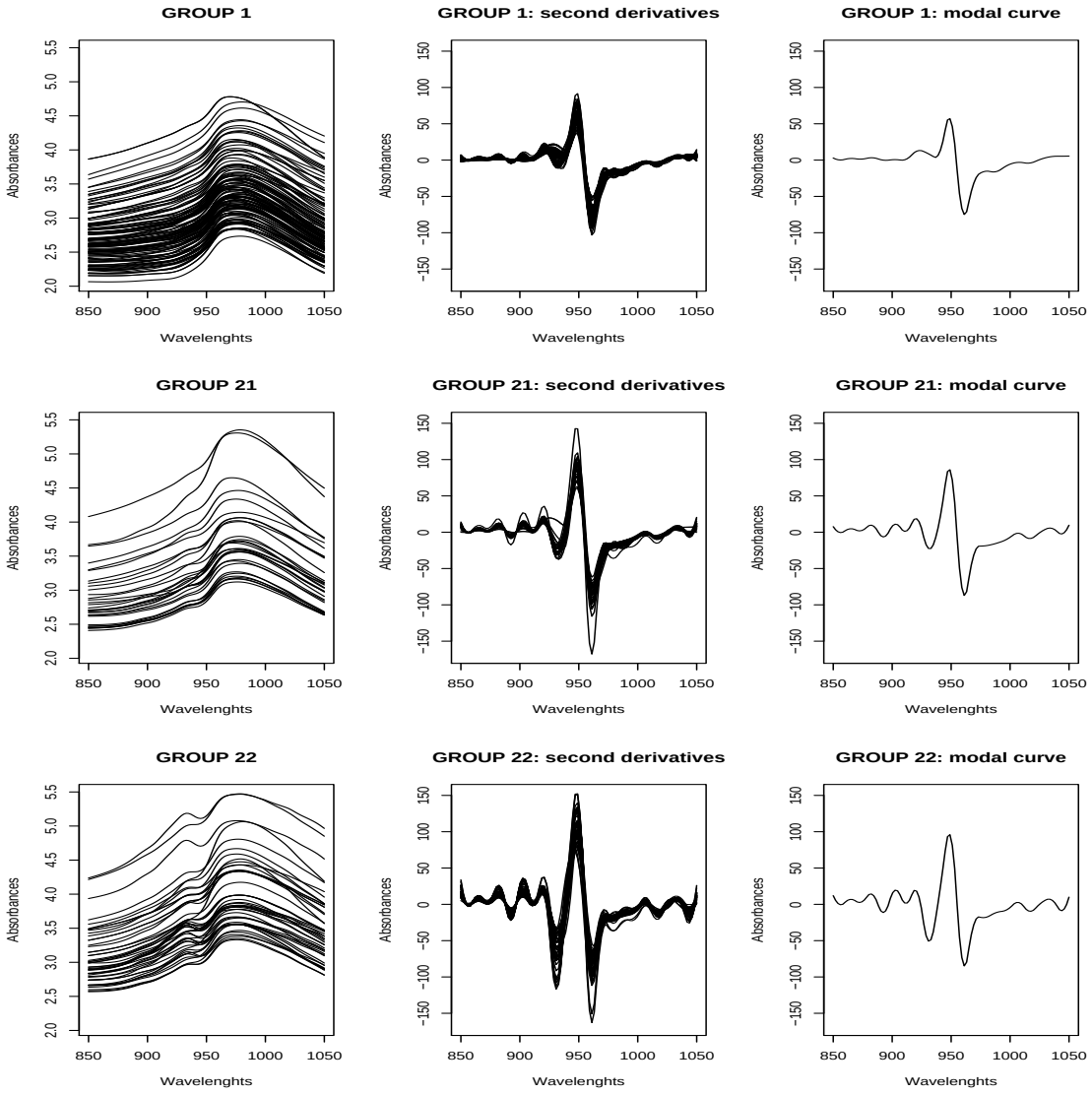

Figure 1: For each group: left column displays the original spectrometric curves, middle column plots their second derivatives and right column displays the corresponding functional mode

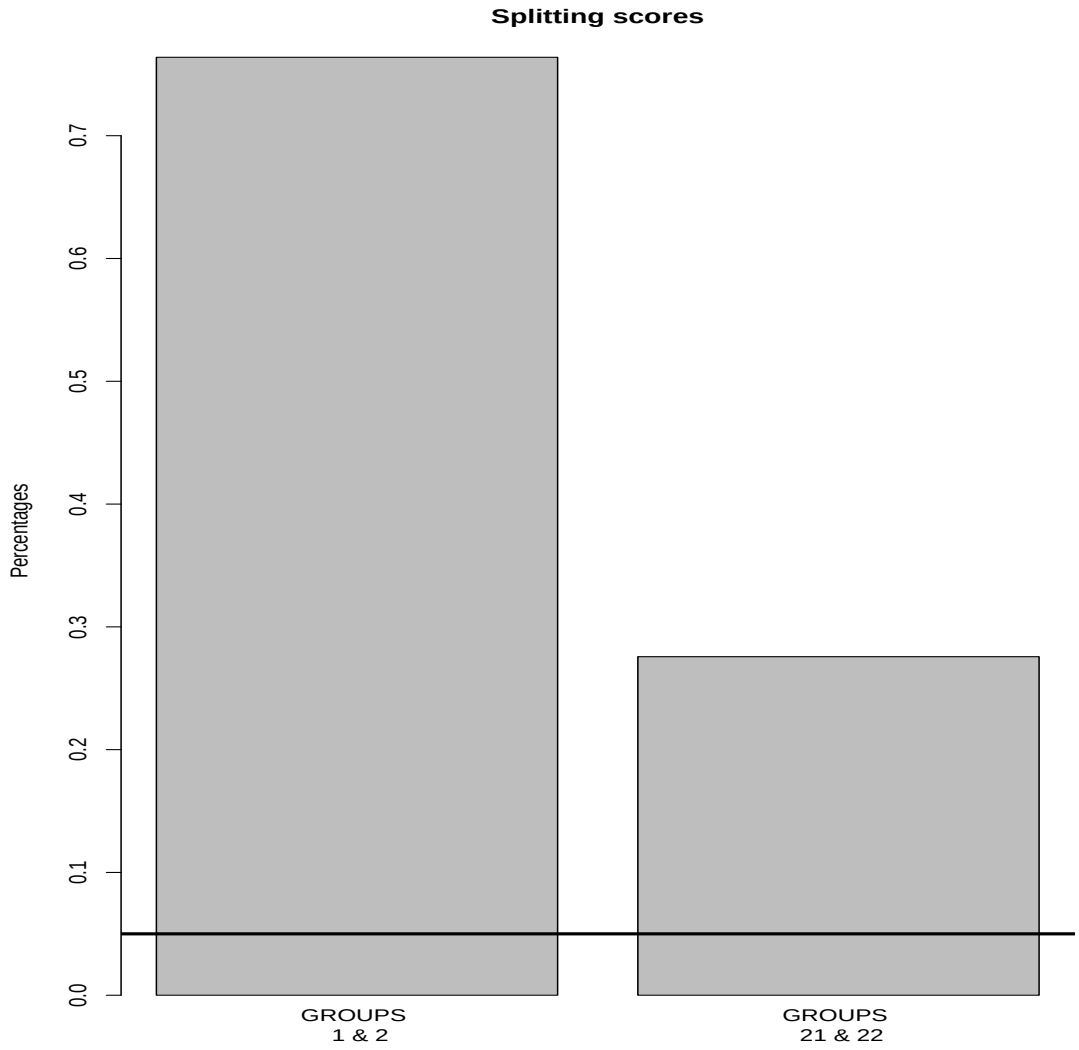

Figure 2: Behavior of splitting score for the spectrometric data; the horizontal line corresponds to the threshold of splitting score