

assignment8

2023-03-31

```
library(MASS)
library(faraway)
library(dplyr)
```

```
##
## 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
```

```
library(pls)
```

```
##
## Attaching package: 'pls'

## The following object is masked from 'package:stats':
##
##      loadings
```

##HW08: Use Lab11_PCR_PLS.R as a guide to perform Partial Least Squares Regression on the fatinmeat data.

```
data(meatspec)
```

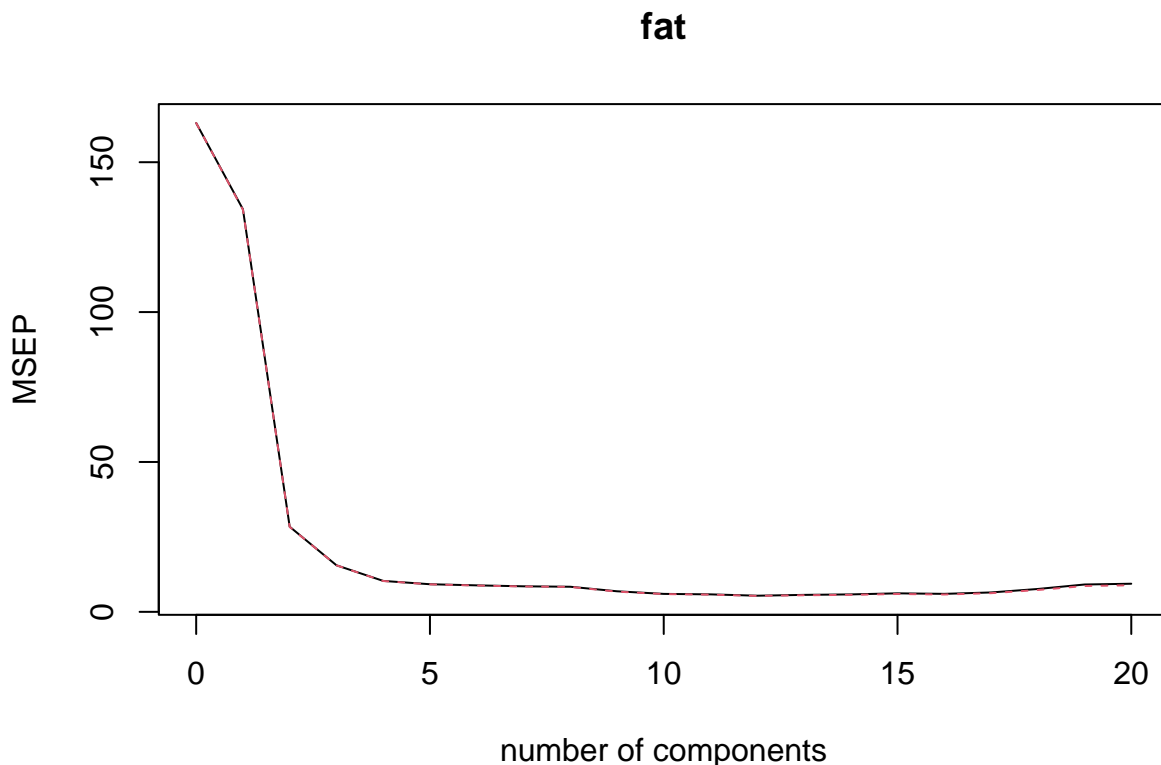
```
meat_data <- meatspec[, 1:50]
meat_data$fat = meatspec$fat
```

#a. Generate a Cross Validation plot using MSE on the Y-axis and the number of components ranging from

```
pls_fit = plsr(fat~., data = meat_data, center=TRUE, scale = TRUE, validation = "CV", ncomp = 20)
summary(pls_fit)
```

```
## Data:      X dimension: 215 50
## Y dimension: 215 1
## Fit method: kernelppls
## Number of components considered: 20
##
## VALIDATION: RMSEP
## Cross-validated using 10 random segments.
##      (Intercept)  1 comps  2 comps  3 comps  4 comps  5 comps  6 comps
## CV           12.77   11.59   5.329   3.941   3.212   3.035   2.972
## adjCV         12.77   11.58   5.323   3.939   3.207   3.029   2.962
##      7 comps  8 comps  9 comps 10 comps 11 comps 12 comps 13 comps
## CV           2.918   2.894   2.619   2.447   2.414   2.320   2.380
## adjCV         2.909   2.884   2.598   2.434   2.396   2.306   2.352
##      14 comps 15 comps 16 comps 17 comps 18 comps 19 comps 20 comps
## CV           2.412   2.484   2.449   2.543   2.762   3.019   3.062
## adjCV         2.381   2.440   2.406   2.496   2.700   2.936   2.977
##
## TRAINING: % variance explained
##      1 comps  2 comps  3 comps  4 comps  5 comps  6 comps  7 comps  8 comps
## X           99.5   99.94   99.99  100.00  100.00  100.00  100.00  100.00
## fat         18.7   83.24   91.05   94.04   94.88   95.27   95.41   95.68
##      9 comps 10 comps 11 comps 12 comps 13 comps 14 comps 15 comps
## X           100.00  100.00   100.0   100.00  100.00  100.00  100.00
## fat         96.53   96.94   97.3    97.59   97.76   97.85   98.02
##      16 comps 17 comps 18 comps 19 comps 20 comps
## X           100.00  100.00   100.00  100.00  100.00
## fat         98.09   98.13   98.16   98.23   98.26
```

```
validationplot(pls_fit, val.type = "MSEP")
```



```
#13 seems to be an ideal choice
```

```
pls_pred = predict(pls_fit, meat_data, ncomp = 13)
```

```
#b. Report the MSEP for the Partial Least Squares Regression based on the number of components you chose
```

```
mean((pls_pred - meat_data$fat)^2)
```

```
## [1] 3.620094
```

```
#MSEP= 3.620094
```