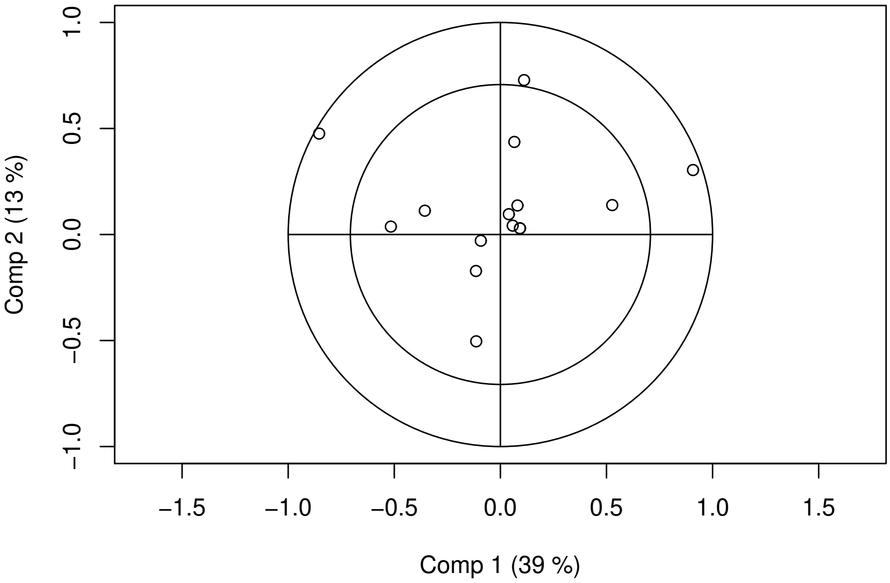
Multivariate Data Analysis
Prof. Dr. Christina Andersson

R Solution Exercise Sheet 1: Partial Least Squares Regression

Computer Problems:

1. (a)
```
> library(pls)
>names(pentatrain)
 [1] "obsnam"  "S1"      "L1"      "P1"      "S2"      "L2"      "P2"
 "S3"      "L3"      "P3"      "S4"      "L4"      "P4"
[14] "S5"      "L5"      "P5"      "log_RAI" "n"
>
```
   (b)
```
gas1 <- plsr(log_RAI ~  S1 +      L1 +      P1 + S2 + L2 +     P2 +
   + P3  +    S4      + L4       + P4 +     S5 +      L5   +    P5, dat
```
   (c)
```
> summary(penta)
Data:   X dimension: 15 15
        Y dimension: 15 1
Fit method: kernelpls
Number of components considered: 14
TRAINING: % variance explained
         1 comps  2 comps  3 comps  4 comps  5 comps  6 comps  7 comps  8 comps  9 c
X          39.38    52.70    66.11    73.97    78.93    90.43    94.56    96.89    9
log_RAI    84.04    96.15    97.87    98.74    98.99    99.06    99.09    99.10    9
         14 comps
X          105.82
log_RAI     99.12
```
   (d) ...

   (e)
```
par(mfrow = c(1,1))
plot(penta, plottype = "correlation")
```



(f) ...

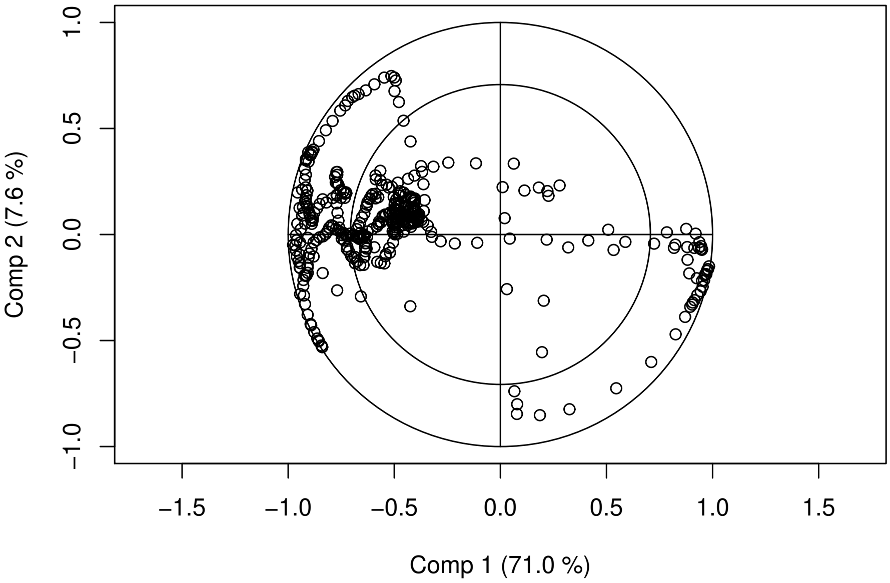
2. (a)
```
> library(pls)
>  names(gasoline)
[1] "octane" "NIR"
```

(b)
```
gas1 <- plsr(octane ~ NIR, ncomp = 10, data = gasoline)
```

(c)
```
> summary(gas1)
Data:   X dimension: 60 401
        Y dimension: 60 1
Fit method: kernelpls
Number of components considered: 10
TRAINING: % variance explained
        1 comps  2 comps  3 comps  4 comps  5 comps  6 comps  7 comps  8 comps  9 co
X         70.97    78.56    86.15    95.40    96.12    96.97    97.32    98.10    98
octane    31.90    94.66    97.71    98.01    98.68    98.93    99.06    99.11    99
```

(d) ...

(e)
```
par(mfrow = c(1,1))
plot(gas1, plottype = "correlation")
```



(f) ...