Multivariate Statistics: Exercise 6

November 20, 2014

Principal component analysis:

We consider the data set *xray.pnm*, which is available from the web page of our exercise.

Install the R package `pixmap`. Read the data into R with

```
x <- read.pnm("xray.pnm")
```

and visualize the object with `plot(x)`. You can see an X-ray false color image with a foreign body in the second finger.

`str(x)` shows the contents of the object. It contains the matrices of the pixels for the color ranges red/green/blue.

Our aim is to compress the information with principal component analysis. Therefore, construct a new object with the same structure as `x`, but where the slots "red", "green", and "blue" are not the original matrices, but reconstructed data using the first k principal components. Thus, apply on each of the 3 matrices the function `prcomp()`, compute loadings and scores, and reconstruct the data with the first k PCs. The resulting new object can be visualized with `plot()`.

- How do you have to reconstruct correctly in order to obtain the same colors as the original image?
- How do you solve the problems that occur when you plot the resulting image?
- Which number of components k do you need to select in order to see the necessary details in the image?
- Compute a measure of information loss when reducing the information to k PCs.
- Compute a "compression factor" informing about the achieved image compression with k PCs.

Save your (successful) R code together with short documentations and interpretations of results in a text file, named as *Familyname7.R*. Send the file as an email attachment to *mehmet.mert@tuwien.ac.at*, at latest Tuesday (18.11)