

# Exercises, STK4040, week 40

September 28, 2007

## Exercise 1

Use last week's data set, and do a 'manual' principal component analysis (PCA); that is, by using the Eigen value decomposition of the covariance matrix.

Compare the results (variances, scores and loadings) with the results from the 'built-in' PCA-function (R: `prcomp()`). Are there any differences, and if so, why?

Does it make any difference if you use the Eigen value decomposition of  $S_u$  or  $S$ ? If so, what is the difference, and which version is 'correct'?

## Exercise 2

Use the data set in the file `dataset_40.txt`. Do a PCA on the data, and interpret as usual. Standardise the data so that all variables have variance 1, and do a PCA on the standardised data. What are the interpretations now? Why is there a difference? Which version gives the 'right' picture? When should we standardise?