

4 SSP 2012 / Assignment 4: PCA & Discriminant

4.1 Reproduction

Reproduce figure 3.3, table 3.1, figures 3.4 and 3.5 from my thesis, Weenink (2006), Adaptive...

4.2 Principal components

1. Generate a two dimensional data set with 1000 elements. The data in the first column are drawn from a normal distribution with mean 0.0 and standard deviation 1, the data in the second column are also drawn from a normal distribution now with mean 0 and standard deviation 0.2. (Use a TableOfReal)
2. Draw a scatter plot of the data (limits -5 and $+5$)
3. Rotate the data 60 degrees counter clock wise. (First cast the TableOfReal to Configuration type. Then use Modify>Rotate... 60 to perform the rotation and then convert back to TableOfReal.)
4. Draw a scatter plot of the data (limits -5 and $+5$) in the same plot as before but with a distinctive color.
5. Perform a principal component analysis on the rotated TableOfReal.
6. Get the fraction of the variance accounted for by component 1 and 2. Are they in accordance with the generated data?

4.3 Discriminant analysis

Create the Pols data (Create TableOfReal (Pols 1973)...) *with and without* formant levels. Use a log transform on the frequencies. Classify with discriminant analysis. What is the difference in percentage correct with and without levels?