Homework 7: Principal Component Analysis (PCA) and Independent Component Analysis (ICA)

[Points: 12.5, Issued: 2008/05/23, Deadline: 2008/06/13, Tutor: Johannes Schweighofer; Info-hour: 2008/05/30, 15:30-16:30, HS i11; Einsichtnahme: 2008/06/20, 15:30-16:30, HS i11;

Please read the tutorial about PCA, ICA, Blind Source Separation carefully before you start the assignment. Accompanying MATLAB programs and data can be downloaded from the homepage. Do not forget to add your name and “Matr.Nr.”. Produce all the plots in the report. Label the axis of your plots. Give explanations to your plots. Provide explanations to your plots and answers. Add to the end of the report the MATLAB script you programmed for producing all the results.

1. Problem: Generate 3-dimensional uniformly distributed data with $N = 2000$ samples using the command `rand()`. Multiply the Data using the following mixing matrix

$$
A = \begin{bmatrix}
2 & 3 & 1 \\
2 & 1 & 0.5 \\
1 & 1 & 1
\end{bmatrix}.
$$

(1)

Visualize the data in the 3-dimensional domain using `plot3()`. Perform the PCA according to Section 3 of the tutorial. For this use commands `cov()` and `eig()`. Once you have determined the principal components draw them into the plot you produced previously using the command `line()`. Transform the data into the first two principal components and visualize them with the command `plot()`. Interpret your results. What happens when you perform PCA? Are the three dimensions of the data (statistically) independent after transformation?

2. Problem: Generate 4 signals using the following commands:

```matlab
N=500
v=[0:N-1];
s=[];
s(1,:)=sin(v/2); % sinusoid
s(2,:)=((rem(v,23)-11)/9).^5; % funny curve
s(3,:)=((rem(v,27)-13)/9); % saw-tooth
s(4,:)=rand(1,N); % random noise
```

Normalize the signals $s$ so that they have zero mean and unit variance by using the command `mean()` and `std()`. Visualize the four signals. Mix the 4 signals by an arbitrary mixing Matrix $A$ (`rand(4)`) to get the sensor signals received at the 4 microphones ($x = As$). Visualize the generated and the mixed signals. Now apply ICA to recover the original source signals $S$ by using the provided command `fastica()`. Plot the resulting signals after applying ICA. Interpret your results. What happens when you perform ICA? Can you establish the original ordering and sign of the signals after applying ICA? Why? Is ICA working for two Gaussian distributed signals $s_i$? Why?

1 mailto:pernkopf@tugraz.at
2 mailto:johannes.schweighofer@student.tugraz.at
3. Problem: Apply ICA similarly as in the previous task using two sound files. You can use the command `wavread()` to load the wav files `sourceX.wav`. Select arbitrarily two different sound files for the experiments. You can play the files using `sound()`. Again plot the loaded, mixed and unmixed signals. Is ICA separating the mixed signals well? Listen to the unmixed signals after applying ICA.

4. Problem: Apply PCA to the previous task using two sound files. Is PCA capable to separate the mixed signal again. Listen and plot the PCA transformed signals? Explain your observations.