Machine Intelligence II **Exercise Sheet 6**
SS 10, Prof. Obermayer due to 17.6.2010

Independent Component Analysis

In this exercise we will implement the Infomax Principle for Independent Component Analysis (ICA) and apply *natural gradient* learning.

Additional Material: Download the data files `sounds.zip` from the ISIS platform.

Write a program that proceeds as follows:

Initialization

(2 points)

- Load the sound files `sound1.dat` and `sound2.dat` (packed in `sounds.zip`). Each of the $N = 2$ sources is sampled at at 8192 Hz and contains $p = 18000$ samples. In Matlab you can use `soundsc` to play them.
- Create a random $N \times N$ mixing matrix $\mathbf{A}$ and mix the sources: $\mathbf{x} = \mathbf{A}\mathbf{s}$
- Permute the columns of the $N \times p$ data matrix $\mathbf{x}$ randomly.
- Calculate the correlations between the sources and the mixtures: $\rho_{\mathbf{s},\mathbf{x}} = \frac{cov(\mathbf{s},\mathbf{x})}{\sigma_{\mathbf{s}}\sigma_{\mathbf{x}}}$
  In Matlab you can use `corr` or `corrcoef`.
- Center the data to zero mean.
- Initialize the unmixing matrix $\mathbf{W}$ with random values.

Optimization

(4 points)

Calculate the *natural gradient* and update matrix $\mathbf{W}$ applying an *online* learning procedure. Use the logistic function for the transformation $\hat{f}$ and choose a suitable learning rate η.

Hint: Implement the matrix formulation of the algorithm. This should reduce your code for this part to one loop (over the samples) and a few lines.

Results

(2 points)

- The recovered signals (estimated sources) are given by: $\hat{\mathbf{s}} = \mathbf{W}\mathbf{x}$
  Play and plot the original sounds, the mixed sources (before and after permutation) and the recovered signals.
- Calculate the correlations (as above) between the true sources and the estimations.

Josef Ladenbauer: `jl@cs.tu-berlin.de`