Machine Learning	Dr. Mario Fritz and Dr. Björn Andres
Exercise Sheet 12 - 02.02.2015	due: 09.02.2015

Exercise 27 - PCA & ICA

This exercise gives you hands-on experience about PCA and ICA.

- **(3 Points) PCA** Complete the function `PCA_diy.m` where you input the raw data `X` and specify the number of reduced dimension `d` and output the projection on the reduced dimension space. In the lecture, two points of view have been described for PCA, i.e.
  - the principal `k`-components span the `k`-dimensional affine subspace which yields the best approximation of the data
  - the subspace spanned by the first `k` princial components contains "most" of the variance in the data

  You can implement **either** of them in your code.
- **(3 Points) Kernel PCA** Complete the function `kPCA_diy.m` with Gaussian kernel with the same input and output format as in `PCA_diy.m`.
- **(1 Points)** Now apply your code (PCA & kPCA) to the provided dataset `pca_data.mat`, visualize your results by running the script `ex12_pca.m` and observe the results. (You don't need to submit any plots on paper!)
- **(3 Points) ICA** Complete the function `ICA_diy.m`. The basic steps have been explained in the lecture "Dimensionality Reduction" page 37, additional information can be found in the tutorial listed in the hints. The function gets input from raw data `X` and specified number of independent components `d` and outputs the obtained components in `Y`. Note the data is in column-wise format, i.e. each column represents a observation.
- **(2 Points) Blind Source Separation** One popular application of ICA is doing Blind Source Separation, i.e. separating a set of source signals from a set of mixed signals without the aid of information (or with very little information) about the source signals or the mixing process `http://en.wikipedia.org/wiki/Blind_signal_separation`. `ica_data.mat` provides you a toy dataset where each row is a mixed signal from two different source signals. Now apply your ICA code to separate the source signals by running the script `ex12_ICA.m`. (You don't need to submit any plots on paper!)

Hints:

- For kernel PCA, you can find more information from `http://www.eecs.berkeley.edu/~wainwrig/stat241b/scholkopf_kernel.pdf`
- For ICA, you can find more information from `http://www.stat.ucla.edu/~yuille/courses/Stat161-261-Spring14/HyvO00-icatut.pdf`

Submission:

- Create **one** zip-file containing the m-files (`PCA_diy.m`, `kPCA_diy.m`, `ex12_PCA.m`, `ICA_diy.m` and `ex12_ICA.m`) and send the file to your tutor. The filename has to follow the following convention: `[group:A,B,C]_[matrikel numbers separated by underscore]_ex[nr].[extension]` e.g. if you are in group B and your team members have matrikelnumbers 3503239, 3028258 and the current exercise number is 14 then the filename reads: `B_3503239_3028258_ex14.zip`.