FastICA & the independent components of image patches

This problem sheet illustrates how to use approximations of *negentropy* (as implemented in the *FastICA* algorithm) to separate mixed toy signals. The second exercise applies this approach to find the independent features of natural images. You can use an available FastICA implementation (fastICA is implemented in scikit-learn and in further toolboxes, cf. http://research.ics.aalto.fi/ica/fastica/).

8.1 Toy Signal Separation (4 points)

Generate three signals as row vectors given at time points $t = 0,\ 0.05, \ldots, 50$ by

$$\begin{aligned} s_1(t) &= 4\ \sin(t-3) \\ s_2(t) &= t + 5\ \bmod 10 \\ s_3(t) &= \begin{cases} -14 & \text{if } \cos(2t) > 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

(a) Mix the signals to get $\mathbf{x} = \mathbf{A}\mathbf{s}$ where

$$\mathbf{A} = \begin{pmatrix} 2 & -3 & -4 \\ 7 & 5 & 1 \\ -4 & 7 & 5 \end{pmatrix}$$

(b) Whiten and separate the mixed (observed) signals $\mathbf{x}$ using `fastICA`.

(c) Plot the original source signals, mixtures, whitened mixtures, and unmixed signals.

(d) Repeat the same analysis using a different contrast function `G`.

(e) Repeat the same analysis using a matrix $\mathbf{A}$ which is closer to singular than the matrix above.

(f) Repeat the same analysis using an additive zero-mean Gaussian noise $\mathbf{n}$ on top of the mixture, i.e., $\mathbf{x} = \mathbf{A}\mathbf{s} + \mathbf{n}$. Try different variances of $\mathbf{n}$ (try first small variances, in another run a larger one etc $\rightsquigarrow$ how robust is the method?).

8.2 fastICA vs. Infomax (2 points)

Apply fastICA to the two soundfiles data set (once again) of problem sheet 6 and compare runtime and robustness w.r.t. the mixing matrix $\mathbf{A}$ with the Infomax-based ICA-algorithm. Use the following setup for the latter algorithm: natural gradient, Bell-Sejnowski amplitude normalization, learning rate schedule $\eta_0 = 0.01$, $\eta_{t+1} = 0.9999\eta_t$.

8.3 ICA on Image Patches (4 points)

The file `imgpca.zip` (used also in exercise sheets 2 and 3) contains three categories of images: *nature*, *buildings*, and *text* (prefixes `n,b,t`). For each category:

(a) Sample P patches of $\sqrt{N} \times \sqrt{N}$ pixels from all images of this category and rearrange each sample to a column vector. Choose number and size of the patches according to your computing resources. Recommended are $P \geq 20000$ and $N \geq 144$.

(b) Calculate the independent features of the image patches (these are the columns of mixing matrix $\mathbf{A}$). Use a `fastICA` toolbox to compute this matrix:

- Let `fastica` perform PCA and whitening of the data.
- Use the contrast function $G(\hat{s}) = \frac{1}{a} \log \cosh(a\hat{s})$ with $a = 1$.

(c) Show the independent features as (grayscale) image patches by rearranging the vectors into $\sqrt{N} \times \sqrt{N}$ matrices and compare the results for the different categories. Order the independent features by decreasing negentropy, (such that the first feature has largest (approximated) negentropy etc).

(d) Perform PCA on the same set of patches, plot the the principal components (ordered by decreasing eigenvalue) as in (c) and compare them with the independent features.

Total points: 10