T-61.5030 Advanced course in neural computing Karhunen, Peltonen

Exercise 6, Nov. 2, 2006

1. Two independent sources s_1 and s_2, both uniformly distributed on $[-1, 1]$, are mixed either with $\mathbf{x} = \mathbf{A_1 s}$ or $\mathbf{x} = \mathbf{A_2 s}$, where

$$\mathbf{A_1} = \begin{pmatrix} 2 & 1 \\ 2 & -1 \end{pmatrix} \text{ and } \mathbf{A_2} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} / \sqrt{2}.$$

   (a) Draw images about how the linear mappings $\mathbf{A_1}$ and $\mathbf{A_2}$ change the coordinate system.

   (b) Compute the covariance matrices of the mixtures.

   (c) Whiten the mixtures with PCA and visualize the new coordinate systems.

   (d) Consider an orthogonal rotation matrix

$$\mathbf{W} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix},$$

   where θ is the angle of rotation. Let u_1 and u_2 be the components of the whitened mixtures after the rotation $\mathbf{W}$. Check that for *any* rotation angle θ, the components u_1 and u_2 remain uncorrelated and have constant variance equal to 1. So, decorrelation or variance maximization will not help in finding the correct rotation $\mathbf{W}$.

   (e) Plot the kurtosis of u_1 and u_2 as functions of θ. How can the kurtosis be used for finding the rotation $\mathbf{W}$.

   (f) Can some other measures be used for finding the rotation?

2. Show that

$$I(X;Y) = D_{p(X,Y)||p(X)p(Y)} ,$$

   that is, mutual information is equal to the Kullback-Leibler divergence of the joint distribution from the "corresponding" factored distribution.

3. Show that for a transformation $\mathbf{y} = \mathbf{Wx}$ there exists a simple connection between mutual information I and negentropy J assuming that the Y_i are uncorrelated and have unit variance:

$$I(Y_1, \ldots, Y_n) = C - \sum_{i=1}^{n} J(Y_i); ,$$

   where C is a constant. (Notation: y_i is the value of the random variable Y_i.) What does this connection imply for ICA computation?