Helsinki University of Technology
Lab of Computer and Information Science

Tik-61.184 Special Course in Information Technology
autumn 2000

Exercises, Session 3

1. Consider the situation where we have a set of L-dimensional observation vectors $\mathbf{x}$ which take on values within the range 0 and 1 i.e. $\mathbf{x} \in [0, 1]$ and satisfies the constraint $\sum_{i=1}^{L} x_i = 1$. We wish to perform an ICA transformation on these variables such that $\mathbf{y} = \mathbf{W}\mathbf{x}$ and the values of the D-dimensional transformed variable $\mathbf{y}$ also satisfy the restrictions $\mathbf{y} \in [0, 1]$ and $\sum_{j=1}^{D} y_j = 1$.

   (a) What restrictions will require to be imposed on the $L \times D$ matrix $\mathbf{W}$ to satisfy the above conditions.? Hint: Think of the range of allowable values of each W_{ij} and the column values of the matrix.

   (b) In devising an ICA routine which will satisfy the identified restrictions on the matrix $\mathbf{W}$ a set of Lagrangian contstraints can be set. An alternative method consists of introducing dummy variables and a nonlinear mapping such that for $\theta_{ij} \in \mathcal{R}$ a mapping $\phi(\theta_{ij}) \to W_{ij}$ exists which satisfies the constraints of (a). Define such a mapping.

   (c) Briefly outline TWO possible applications of such an ICA routine which satisfies either or both of the given constraints.

2. The infomax algorithm was shown to provide sparse localised and oriented basis vectors for natural images when the logistic function is used as the network nonlinearity.

   (a) What is the form of the implicit prior probability on the coefficient values of the image synthesis model that the logistic nonlinearity imposes?

   (b) If a double exponential prior, $p(s) \propto exp(-\beta|s|)$, on the coefficients is assumed give the associated nonlinear function which would appear in the infomax iterative update.

   (c) When using appropriate nonlinear functions ICA methods search for both sparsity **and** independence. Discuss the relative importance of these two characteristics in finding the natural image basis functions given in the Session 3 papers.

3. The overcomplete basis set method detailed in 'Sparse coding with an overcomplete basis set: A strategy employed by V1?' could be used to extract **more** sources than there are available observations in the case of blind signal separation. Give a detailed discussion on the practicality of this proposal.