Humboldt-Universität zu Berlin
Institute for Theoretical Biology

Exercise Set 5
Summer 2006, May 22

Computational Neuroscience IV: Analysis of Neural Systems

Dr. R. Kempter, Prof. Dr. A.V.M. Herz

Information Theory

ICA estimation by minimization of mutual information I is equivalent to maximizing the sum of negentropies (or "nongaussianities") J_i of the estimates y_i $(i = 1, \ldots, n)$ of the independent components if the y_i are constrained to be uncorrelated.

Given a random vector $\mathbf{x}$ and an invertible linear transformation $\mathbf{y} = \mathbf{B}\,\mathbf{x}$, mutual information and negentropy differ only by the sign and a constant that does not depend on $\mathbf{B}$,

$$I(y_1, y_2, \ldots, y_n) = \text{const.} \; - \sum_{i=1}^{n} J(y_i) \; . \tag{1}$$

1. Compute the constant in equation (1) when the y_i are constrained to unit variance.

2. If the variances of the y_i are not constrained to unity, how does the constant in equation (1) change?

3. Compute the mutual information for a Gaussian random vector $\mathbf{y}$ with covariance matrix $\mathbf{\Sigma}$.

Infomax Principle for a Network of Neurons

4. Assume that $\mathbf{x} = (x_1, \ldots, x_n)$ is the input to a 'neural network' whose output $\mathbf{y} = (y_1, \ldots, y_n)$ is of the form

$$y_i = \phi_i(\mathbf{b}_i^T \mathbf{x}) + n_i$$

where n_i is Gaussian white noise, ϕ_i is a sigmoidal scalar function, and $\mathbf{b}_i$ is the weight vector of neuron i. One goal of network design or 'learning' is to maximize the entropy $H(\mathbf{y})$ of the output. On the other hand, efficient information transmission in the network requires that the mutual information $I(\mathbf{x}, \mathbf{y})$ of input and output be maximized. Show that both principles are equivalent by employing a similar strategy as in Problem 6 of Exercise Set 4. Compare the result to the likelihood in the noise-free ICA model. (Hint: for an invertible transformation $\mathbf{u} = \boldsymbol{\psi}(\mathbf{v})$ use the entropy transformation $H(\mathbf{u}) = H(\mathbf{v}) + E\{\log|\det \frac{\partial}{\partial \mathbf{v}} \boldsymbol{\psi}(\mathbf{v})|\}$ where $\frac{\partial}{\partial \mathbf{v}} \boldsymbol{\psi}(\mathbf{v})$ is the Jacobian matrix of $\boldsymbol{\psi}$; also use $\boldsymbol{\phi}(\mathbf{v}) = [\phi_1(v_1), \ldots, \phi_n(v_n)]^T$)

Time Filtering

Basic data preprocessing such as centering, whitening, and PCA-based reduction of the dimension of the data can improve and simplify ICA. Time filtering is not necessary in theory, but often very useful in practice when the observed random variables are time series and the sample index $t = 1, \ldots, T$ of the data $\mathbf{x}(t) = [x_1(t), \ldots, x_n(t)]^T$ is a time index. The time-filtered data then is

$$x_i^*(t) = \sum_{t'=1}^{T} x_i(t')\, m_{t',t} \quad \text{for } i = 1, \ldots, n$$

where $m_{t',t}$ is an element of the $T \times T$ filter matrix $\mathbf{M}$. Note that x_i and x_j are not mixed if $i \neq j$.

5. We denote by $\mathbf{X}$ the $n \times T$ matrix that contains the observations, $\mathbf{X} = [\mathbf{x}(1), \ldots, \mathbf{x}(T)]$, and, similarly, the sources $\mathbf{S} = [\mathbf{s}(1), \ldots, \mathbf{s}(T)]$. The ICA model then is $\mathbf{X} = \mathbf{AS}$.

a) Indicate the structure of $\mathbf{M}$ for low-pass filtering or smoothing, which means that every sample point $x_i(t)$ is replaced by a weighted average over adjacent sample points. Explicitly take the borders $t = 0$ and $t = T$ into account.

b) How does $\mathbf{M}$ look like for high-pass filtering, where every sample point is replaced by the difference between two neighboring points?

c) Show that the mixing matrix $\mathbf{A}$ for the time-filtered model remains unchanged, $\mathbf{X}^* = \mathbf{AS}^*$.

d) In contrast to time-filtering, why does whitening change the ICA model, although it can also be denoted by a matrix multiplication?

Dr. R. Kempter, phone 2093-8925, room 2315, `r.kempter(AT)biologie.hu-berlin.de`
Prof. Dr. A.V.M. Herz, phone 2093-9112, room 2325, `a.herz(AT)biologie.hu-berlin.de`

Problems handed out on Monday, 22.05.2006. Solutions to be handed in by Monday, 29.05.2006, 12^{15} pm. Discussion and presentation of the problems on Friday, 02.06.2006, $8^{30} - 10^{00}$ am in front of room 2317 I-W (Zwischengeschoß).