Assignment 6

Probabilistic and Unsupervised Learning

Maneesh Sahani

Due: Thursday Dec 17, 2009

1. [40 marks] Expectation-Propagation


(a) First, write down the log-joint probability log(p(\textbf{s}, \textbf{x})). Rearrange the terms to form a sum of log-factors \sum_i \log f_i(\textbf{s}), where each \textit{f}_i collects any terms that depend on \textit{s}_i alone, as well as a “symmetric” share of the cross-terms. Remember that \textit{s}_i^2 = \textit{s}_i.

(b) Next, find the expected value (with respect to \textit{q}_{jn} for \textit{j} \neq \textit{i}) that minimises the partial KL:

\[ \text{KL} \left[ \prod_{\textit{j} \neq \textit{i}} \textit{q}_{jn}(\textit{s}_j^{(n)}) \textit{f}_i(\textbf{s}^{(n)}) \right| \prod_{\textit{j} \neq \textit{i}} \textit{q}_{jn}(\textit{s}_j^{(n)}) \textit{q}_{in}(\textit{s}_i^{(n)}) \]  

Here \textit{q}_{jn}(\textit{s}_j^{(n)}) is the \textit{j}th factor in \textit{q}_n(\textbf{s}^{(n)}). It might help to reparameterise in terms of the natural parameter \textit{\rho}_{jn} instead of \textit{\lambda}_{jn} as above.

(c) Finally, find the update rule for \textit{\rho}_{in}. It might to differentiate the expected value derived above with respect to \textit{s}_i.

2. [10 marks] Describe a Bayesian method for selecting \textit{K}, the number of hidden binary variables in this model. Does your method pose any computational difficulties and if so how would you tackle them?

3. [Bonus: 50 marks] Implement the EP algorithm you derived above, and compare your results to those of the variational mean-field algorithm of Assignment 4.