Assignment 2: Latent Variable Models

Unsupervised Learning

Zoubin Ghahramani

NEW Due: Thurs Oct 23, 2003

Note: The Matrix Inversion Lemma

$$(A + XBX^\top)^{-1} = A^{-1} - A^{-1}X(B^{-1} + X^\top A^{-1} X)^{-1} X^\top A^{-1}$$

is a useful tool to know, and may be useful for some of these questions.

1. Posterior over Factors in Factor Analysis
In Factor Analysis:

$$p(\mathbf{x}) = N(0, I)$$

$$p(\mathbf{y}|\mathbf{x}) = N(\Lambda\mathbf{x}, \Psi)$$

Derive the expression for the mean and covariance of $p(\mathbf{x}|\mathbf{y})$. Hint: write out the joint distribution $p(\mathbf{x}, \mathbf{y})$ and treat $\mathbf{y}$ as a constant.

2. Principal Components Analysis
In Probabilistic Principal Components Analysis

$$p(\mathbf{x}) = N(0, I)$$

$$p(\mathbf{y}|\mathbf{x}) = N(\Lambda\mathbf{x}, \sigma^2 I)$$

and the principal components are assumed to be orthonormal: $\Lambda^\top\Lambda = I$. Derive the mean and covariance of $p(\mathbf{x}|\mathbf{y})$ in the PCA limit, $\sigma^2 \to 0$.

3. The k-means algorithm
Read Chapter 20 in David MacKay's book (pp 284-292) paying attention to why the k-means algorithm often does not work well. Implement the k-means algorithm in Matlab. Run your algorithm on two different two-dimensional data sets similar to the ones in MacKay's book showing situations where the k-means algorithm does not find the clusters one would hope a clustering algorithm would. Hand in: a printout of the Matlab code, plots of the two data sets with the means found by k-means clearly indicated.

4. Determining the Number of Gaussians in a Mixture Model
Assuming that a data set was produced from a mixture of axis-aligned Gaussians, specify a method for deciding how many Gaussians should be fitted to the data set. Try to specify your method in sufficient detail so that one of your classmates could follow it mechanically, but do not get carried away. The specification should be less than one page.