HMM with mixture model emissions

A common modification of the hidden Markov model involves using mixture models for the emission probabilities \( p(y_t|q_t) \), where \( q_t \) refers to the state for time \( t \) and \( y_t \) to the observation for time \( t \).

Suppose that \( y_t \in \mathbb{R}^n \) and that the emission distribution is given by a mixture of Gaussians for each value of the state. To be concrete, suppose that the \( q_t \) can take \( K \) discrete states and each mixture has \( M \) components. Then,

\[
p(y_t | q_t) = \sum_{j=1}^{M} b_{q_t j} \left( \frac{1}{(2\pi)^{n/2} |\Sigma_{q_t j}|^{1/2}} \exp \left\{ -\frac{1}{2} (y_t - \mu_{q_t j})^T \Sigma_{q_t j}^{-1} (y_t - \mu_{q_t j}) \right\} \right)
\]

where \( b_j \in [0, 1]^M \) denotes the mixing weights for state \( i \) (\( \sum_{j=1}^{M} b_{ij} = 1 \) for \( i = 1, \ldots, K \)), \( \mu_{ij} \in \mathbb{R}^n \) and \( \Sigma_{ij} \in \mathbb{R}^{n \times n} \).

Let \( \pi \in \mathbb{R}^K \) be the probability distribution for the initial state \( q_0 \), \( A \in \mathbb{R}^{K \times K} \) the transition matrix of the \( q_t \)'s. In this problem you will derive an EM algorithm for learning the parameters \( \{b_{ij}, \mu_{ij}, \Sigma_{ij}\} \) and \( A, \pi \).

1. The EM algorithm is substantially simpler if you introduce auxiliary variables \( z_t \in \{1, \ldots, M\} \) denoting which mixture component the \( t \)’th observation is drawn from.

Draw the graphical model for this modified HMM, identifying clearly the additional latent variables that are needed.

2. Write the expected complete log likelihood for the model and identify the expectations that you need to compute in the E step. Show all steps of your derivation.


   Hint: Reduce the inference problem to something you know how to do, such as sum-product belief propagation in tree-structured pairwise MRFs.

4. Write down the equations that implement the M step.