T.61.5140 Machine Learning: Advanced Probablistic Methods
Hollmén, Raiko (Spring 2008)
Problem session, 28th of March, 2008
http://www.cis.hut.fi/Opinnot/T-61.5140/

1. Given a hidden Markov model (HMM, page 610) and observations $\mathbf{y}_1, \ldots, \mathbf{y}_{t-1}$, show that the predictive distribution of the observations $\mathbf{y}_t$ at time point t follows a mixture distribution.

2. Show how a second-order Markov chain (page 608) of 3 symbols can be transformed to a hidden Markov model with 9 states and 3 symbols.

3. Let us consider a HMM with a discrete hidden variable z with 6 states and a Gaussian observation (emission) probability density function. The dimension of the data vectors $\mathbf{x}_1, \ldots, \mathbf{x}_T$ is 5 and the covariance function of the Gaussian distribution is diagonal. (a) Quantify the number of parameters in the model, (b) write the joint probability density, (c) and write the Q-function of the EM-algorithm $Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{\text{old}})$ (page 440). Assume that the E-step is done, that is, $\gamma(z_t) = P(z_t \mid \mathbf{X}, \boldsymbol{\theta}^{\text{old}})$ and $\xi(z_{t-1}, z_t) = P(z_{t-1}, z_t \mid \mathbf{X}, \boldsymbol{\theta}^{\text{old}})$ are given.

4. In the setting of Problem 3, (a) derive the M-step for the Gaussian means μ_{ik}, where $i = 1 \ldots 6$ denotes the state and $k = 1 \ldots 5$ denotes the data dimension. (b) Derive the M-step for updating the 6×6 transition matrix $\mathbf{A}$.