
Home Work 3

Due Date: April 21st

Exercise

In this homework you need to encode the EM algorithm to estimate the transition **A** and observation **B** matrices for the provided sequence.

First encode the forward-backward procedure, assuming **A** and **B** are known. This function takes as input the sequence x_1, \dots, x_T (file seq.txt) that takes the values 1 to 10, **A** and **B**. The output is the posterior probability for z_t . For example if T is 1000 and **A** is 3 by 3. The output is a matrix of 3 by 1000 in which each column adds up to 1.

Second encode the maximum likelihood estimate for **A** and **B**. The input is the sequences x_1, \dots, x_T and z_1, \dots, z_T . to obtain the sequence of the hidden variables, do not take the maximum, but sample from the posterior probability.

Repeat between both steps until convergence. This procedure is not convex, so it might return a local minima. To simplify finding the right solution the true transition matrix is:

$$\mathbf{A} = \begin{bmatrix} 3/4 & 1/4 & 0 \\ 1/4 & 1/2 & 1/4 \\ 0 & 1/4 & 3/4 \end{bmatrix}$$

- The code and instructions in how to run it in a mac.
- The estimates for **A** and **B**.
- A 200-word paragraph (or less) with your conclusions.