Problem Set 4

CS 6347

Due: 4/6/2017 by 11:59pm

Note: all answers should be accompanied by explanations for full credit. Late homeworks cannot be accepted. All submitted code **MUST** compile/run.

Problem 1: Maximum Likelihood for Colorings (100 pts)

For this problem, we will use the same factorization as we have in past assignments. However, the weights will now be considered parameters of the model that need to be learned from samples.

1. Use the belief propagation algorithm that you wrote on Problem Set 2 to perform (approximate) maximum likelihood estimation for the coloring counting problem. Your solution should be written as a MATLAB function that takes as input an $n \times n$ matrix A corresponding to the adjacency matrix of the graph G and `samples` which is an $n \times m$ binary matrix whose t^{th} column corresponds to the t^{th} sample. The output should be the matrix of weights w corresponding to the MLE parameters learned for each variable.

   ```
   function w = colormle(A, samples)
   ```

   Your algorithm should return the correct maximum-likelihood parameters when the graph is a tree and approximate parameters otherwise. You can test your algorithm by generating a large number of samples from a tree-structured model using the Gibbs sampler that you wrote on Problem Set 3.

2. Suppose that you are given a node $i \in V$ such that i is a latent variable in the model. Given m samples of the observed variables in $V \setminus \{i\}$, what is the log-likelihood as a function of the weights? Perform MLE using the EM algorithm. Your solution should be written as a MATLAB function that takes as input an $n \times n$ matrix A corresponding to the adjacency matrix of a graph G, the vertex $i \in V$, and `samples` which is an $n \times m$ binary matrix whose t^{th} column corresponds to the t^{th} sample (note that the component corresponding to i can be any value and it should not be used by your algorithm). The output should be the matrix of weights learned using the EM algorithm.

   ```
   function w = colorem(A, i, samples)
   ```