

Assignment 2: Parameter Estimation and Approximate Inference

CS786 – Spring 2012

Out: Thursday, June 21, 2012
Due: Monday, July 9, 2012, at midnight

This assignment consists entirely of programming questions to be done in the Java programming language. You will implement some basic functions for parameter estimation and approximate inference with Bayesian networks and Markov networks. Submission is done electronically via Marmoset. The Marmoset server will automatically compile your code and run various tests to verify its correctness. Your grade will be the number of public and release tests passed by your code.

A skeleton jar file that includes the definitions of several Java classes and the signature of the methods that you need to implement will be posted on the course website shortly. Your job is to fill in several methods for parameter estimation and approximate inference:

1. **estimatedBayesNet = supervisedMLEstimation(bayesNet, data):** function that estimates the conditional probability tables of a Bayesian network from complete data by maximum likelihood. The input Bayesian network consists of a graph without conditional probability distributions and the output Bayesian network is the same graph with conditional probability distributions.
2. **estimatedBayesNet = expectationMaximization(bayesNet, data):** function that estimates the conditional probability tables of a Bayesian network from incomplete data by expectation maximization (EM). EM performs semi-supervised learning based on incomplete data. The data is incomplete in the sense that the values of some attributes are missing in some records. The input Bayesian network consists of a graph without conditional probability distributions and the output Bayesian network is the same graph with conditional probability distributions.
3. **approximateMarginals = pgm.loopyBeliefPropagation(evidence):** compute the approximate marginal of each non evidence variable by loopy belief propagation (a.k.a. cluster graph propagation). Define a cluster node for each variable and each factor in the Bayes net or Markov net. Return the factors associated with the cluster nodes that consist of single variables. These factors are the approximate marginals of the non evidence variables.
4. **approximateAnswer = bayesNet.likelihoodWeighting(queryVariables, evidence):** compute the approximate answer to an inference query by likelihood weighting.
5. **approximateAnswer = pgm.gibbsSampling(queryVariables, evidence):** compute the approximate answer to an inference query by Gibbs sampling.