Problem Set 3

CS 6347

Due: 3/12/2016 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: The Approximate MAP Problem Revisited (20 pts)

On the last problem set, you estimated the maximum weight coloring by running the max-product algorithm. By plugging the approximate max-marginals into the MAP LP, you could attempt to estimate the value of the maximum coloring. However, the max-marginals satisfy a max-marginalization condition, not a marginalization condition, so they need not exist in the local polytope (even after renormalization). So, in general, we can't just plug approximate max-marginals into the MAP LP in order to obtain an estimate of the MAP value. If the maximum value is unique for each belief, you can plug it directly into the objective, but what if there are ties?

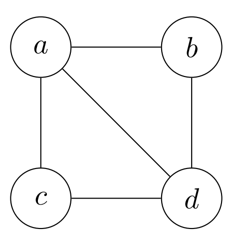
1. Suppose you run max-product to convergence. Let m^* represent the converged messages and b^* represent the converged beliefs (as defined on Problem Set 2). Show that $\max_{x_j} b^*_{ij}(x_i, x_j) \propto b^*_i(x_i)$.

2. Suppose you run sum-product to convergence. Let m^* represent the converged messages and b^* represent the converged beliefs (as defined on Problem Set 2). Show that $\sum_{x_j} b^*_{ij}(x_i, x_j) \propto b^*_i(x_i)$.

3. We can obtain approximate marginals for the approximate MAP problem over the local marginal polytope by using sum-product. Explain how to do this using the fact that $\max_x p(x) = \lim_{T \to 0} \left(\sum_x p(x)^{1/T}\right)^T$.

The real number T is usually called the temperature. More colloquially, we say that max-product is the zero temperature limit of sum-product.

Problem 2: Gibbs Sampling (60 pts)



1. Use the Gibbs sampling algorithm to approximate the marginals of the probability distribution over weighted colorings described in problem 2 of problem set 2. 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, a vector of weights $w \in \mathbb{R}^k$, `burnin` which is the number of burn-in samples, and `its` which is an input that controls the total number of iterations of the Gibbs sampler after burn-in. The output should be a matrix of marginals whose i^{th}, x_i^{th} component is equal to the probability that $i \in V$ is colored with color $x_i \in \{1, \ldots, k\}$.

```
function m = gibbs(A, w, burnin, its)
```

2. Run the Gibbs sampler to estimate the probability that $a \in G$ is colored with color 4 using weights $w_i = i$ for each $i \in \{1, 2, 3, 4\}$ in the graph above. Construct a table of the estimated marginal as a function of `burnin` versus `its` for your implementation where `burnin` and `its` are chosen from the set $\{2^6, 2^{10}, 2^{14}, 2^{18}\}$. Does your answer depend on the initial choice of assignment used in your Gibbs sampling algorithm?

Problem 3: Maximum Likelihood (20 pts)

Given a graph $G = (V, E)$, the joint distribution corresponding to the Ising model is a probability distribution over random variables $x \in \{-1, 1\}^{|V|}$ and is expressed as

$$p(x) = \frac{1}{Z} \prod_{i \in V} \exp(h_i x_i) \prod_{(i,j) \in E} \exp(J_{ij} x_i x_j)$$

for some choice of parameters h and J.

Let G be a complete graph on three nodes. Suppose that we are given the following samples from some unknown Ising model over G: $\{-1, -1, 1\}, \{1, -1, -1\}, \{1, 1, 1\}, \{-1, -1, -1\}, \{1, -1, -1\}$. Assuming that $h_i = 0$ for each vertex and $J_{ij} = J$ is chosen to be the same for each edge, use maximum likelihood estimation to estimate the parameter J from these samples.